UNIVERSITY OF TASMANIA

The role of the *FT* genes in the control of flowering in chickpea

Raúl Ortega Martínez

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy
School of Biological Sciences, University of Tasmania, January 2018

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Authority of Access

This thesis may be made available for loan. Copying and communication of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.

_____

Raúl Ortega Martínez

Date: 29 January 2018

Abstract

The transition from vegetative to reproductive stage is one of the most significant in the life cycle of any plant. Variation in flowering time allows species to colonize new habitats, and from an agricultural point of view is crucial to adapt crops to different environments and maximise yield. This is particularly important in the case of chickpea, which is cultivated in diverse environments in more than 50 countries. In most of these environments, early phenology plays a key role as an adaptation that allows the crop to escape unfavourable conditions. Therefore, flowering time is one of the most extensively studied traits in chickpea and numerous QTL studies have been published on this topic. *FLOWERING LOCUS T (FT)* is a key gene promoting flowering in *Arabidopsis thaliana*, and *FT* homologs are involved in the control of floral transition across plant kingdom, including legumes. This study investigated the putative role of *CaFT* homologues in the genetic control of flowering in chickpea.

The molecular control of flowering time is well understood in model species, particularly Arabidopsis, which is the most suitable model for comparison with temperate legumes in view of its taxonomic position and nature of its flowering responses. This study explored the conservation and position of Arabidopsis flowering-related genes across chickpea genome, and discussed their co-localization with some reported flowering QTLs, focussing in particular on a central portion of chickpea chromosome 3 that has been recurrently associated with flowering in several mapping populations and published studies. The most plausible candidates in this region, belonged to the well-known *CONSTANS-Like* (*COL*) and *FT* gene families and these gene families were therefore characterised.

Three different intra- and inter-specific chickpea populations were used in this study to investigate in more detail the possible identity of the genes underlying the co-localized flowering time QTLs on chromosome 3. QTL analysis and differential expression profiles in the three populations identified a cluster of three *FT* homologs (*FTa1-FTa2-FTc*) as the genes most likely to be responsible for the majority of the phenological difference between wild and cultivated chickpea. In contrast, in the intraspecific population, this locus has a lesser role that was secondary to a major locus in another region of chromosome 3. QTL analysis of shoot architecture traits revealed major loci controlling growth habit (erectness) and branching tendency between *C. arietinum* and *C. reticulatum* located in the same interval

of chromosome 3, indicating a possible pleiotropic role of *FT* genes in control of shoot architecture. The growth habit QTL is likely equivalent to the previously-described growth habit locus *Hg*, suggesting that the *FT* cluster should be considered candidates for this locus.

In view of their potential to influence chickpea phenology and thus its adaptation to different environments, sequence variation of the chickpea *FTa1-FTa2-FTc* cluster was examined in a panel of 96 accessions (94 *C. arietinum* and 2 *C. reticulatum*) through a targeted next-generation sequencing approach. This analysis showed that despite high conservation within the coding regions, the regulatory and intergenic regions are very divergent in the wild and domesticated species. Among domesticated accessions, the intergenic region between *FTa1* and *FTa2* shows the highest level of polymorphism, including the total deletion of the *FTa2* gene and a 753 bp insertion that could be associated with variation in flowering time.

Vernalization response has been a controversial topic in chickpea. Unlike *C. reticulatum* from which it derives, *C. arietinum* has been traditionally considered as a vernalization-insensitive species. However, more recent evidence points to the existence of two distinct vernalization response patterns within cultivated germplasm. This study evaluated the flowering phenotype of two wild and six cultivated chickpea accessions in response to vernalization and photoperiod, and the potential role of the chickpea *FT* homologs in these pathways was investigated through analysis of their expression patterns. The results indicate that a response to vernalization exists in all *C. arietinum* accessions, and suggest that the *FTa1* gene may be particularly important in the signalling and integration of photoperiod and vernalization responses. The convergence of these two pathways on overlapping groups of *FT* genes may explain why a subset of chickpea accessions behave like they are vernalization insensitive under flowering-inductive photoperiods.

Overall, the results obtained in the present study make a significant contribution to the current understanding of regulation of flowering time and growth habit in chickpea, including the molecular basis for a major flowering time locus and potential roles for *FT* homologues in the control of flowering time in both wild and domesticated chickpea.

Acknowledgements

I would like to thank all the people who offered their help and support during the process of creating this thesis. First, I would like to express my sincere gratitude to my supervisors in Hobart, Jim Weller and Valérie Hecht, not only for their excellent guidance and priceless assistance during all my PhD but also for their friendship and extraordinary kindness towards me and my family at every moment.

I am very thankful to the people in the Department of genetic of the University of Cordoba (Spain), in particular to my supervisor Teresa Millan for her encouragement to start this PhD and for sharing her expertise in the physiology of chickpea and management of populations. In the same department, I would like to extend my gratitude to Juan Bautista Alvarez and Carlos Guzman, for introducing me in the academic world and their support at the very beginning, when it is more needed.

I also want to thank Jose Rubio and people at the IFAPA centre in Cordoba for their invaluable contribution growing and phenotyping the populations in the field. Their work was essential for the success of this thesis.

I would like to thank Dr. Tim Sutton (South Australia Research and Development Institute) for generously sharing the chickpea collection used in chapter 6. I am also grateful to Dr. Douglas Cook (United States Department of Agriculture, Agricultural Research Service and Washington State University) for facilitating the molecular markers for CRIL2 population.

My deepest gratitude for my fellows in the Pea Flowering research group: Jackie Vander Schoor, Stephen Ridge, Francess Sussmilch, Bea Contreras, Owen Williams and Vino Rajandran, for their friendship, help and support.

I am grateful to my friends and fellow researchers Carlos, Tito, Tania, Eva, Jaouad, Sebas and Jose. Without them I would have never started this journey. I cannot forget my old friends in Cordoba from the school and neighbourhood, who make me laugh every single day despite the distance and time difference separating us. I am also truly thankful to my friends in Hobart, who opened their houses and hearts and became a second family, particularly John, Elisa, Carmen, Quique and Bea again. I was very lucky to meet such a wonderful group of people and share this time with you.

Thank you to all PhD candidates, professors and staff at the School of Biological Science for creating an awesome environment to grow not only as scientist but also as a person. Special thanks to the people in charge of the controlled environment facilities in Hobart, Michelle Lang and Tracey Winterbottom, for looking after my plants during all my PhD. I would also like to acknowledge Dr. Jules Freeman for its help with the construction of genetic map and QTL analysis, and to Adam Smolenski and Sharee McCammon for their technical support and advice in the laboratory.

I would like to thank the wonderful people sharing my office during this 4 years; Son, Vino, Alvin, Bich and, more recently Andy, for their support, help and infinite patience with me.

I thank the University of Tasmania for providing funding for my scholarship.

Last but not least, I would like to thank my family, specially my parents, for their love, optimism and tireless support during this journey. I would like to dedicate this thesis to the most important persons in my life, my wife Teresa and my son Aníbal. For their love and being the pillars of my life, source of strength and inspiration.

Table of contents

List of Figures

List of Tables

Abbreviations

bp	Nucleotide base pairs
CAPS	Cleaved Amplified Polymorphic Sequences
cDNA	Complementary DNA
CDS	Coding sequence
CO	*CONSTANS*
COL	*CONSTANS-LIKE*
DArT	Diversity Arrays Technology
DFTP	Days from opening of the first flower to formation of first pod
DNA	Deoxyribonucleic acid
DTF	Days to flowering
EST	Expressed Sequence Tags
FT	*FLOWERING LOCUS T*
Gbp	Giga base pair
HRM	High Resolution Melt
ICARDA	International Center for Agricultural Research in the Dry Areas
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
ISSR	Inter-Simple Sequence Repeat
LD	Long day(s)
LG	Linkage Group
lncRNA	Long non-coding RNA
MAFFT	Multiple Alignment using Fast Fourier Transform
Mbp	Mega base pair
mRNA	messenger RNA
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
NCBI	National Center for Biotechnology Information
ncRNA	Non-coding RNA
NFD	Node of flower development
NFI	Node of flower initiation
NPD	Node of pod development
ORF(s)	Open Reading Frame(s)
PCR	Polymerase Chain Reaction

qPCR	Quantitative Polymerase Chain Reaction
QTL	Quantitative Trait Loci
RIL	Recombinant Inbred Line
RIP	Recombinant Inbred Population
RNA	Ribonucleic acid
RT-PCR	Real-Time Polymerase Chain Reaction
SD	Short day(s)
SDW	Sterile Milli-Q water
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
STMS	Sequence Tagged Microsatellite Markers
UTR	Untranslated region

Chapter 1. General introduction

1.1 General botany.

Chickpea (*Cicer arietinum* L.) is a diploid plant with *2n=2x=16* chromosomes and a genome size of 738 Mbp (Varshney et al. 2013c). It is an annual, self-pollinated species belonging to the family *Fabaceae*, subfamily *Papilionoideae*, its own tribe *Cicereae* Alef, and the genus *Cicer*. This genus includes 44 species, of which 9 are annual and 35 perennials (Table 1.1) (Van Der Maesen et al. 2007; Van der Maesen 1987), and *C. arietinum* is the only cultivated species.

Table 1.1 List of the species belonging to the genus *Cicer*

Subgenus Viciastrum			**Subgenus Pseudononis**
Section Acanthocicer	**Section Polycicer**		**Section Monocicer**
C. acanthophyllum Boriss.	*C. analolicum* Alef.	*C. korshinskyi* Lincz.	*C. arietinum* L.
C. incanum Korotk.	*C. allanlicum* Coss. ex Maire	*C. microphyllum* Benth.	*C. bijugum* K.H. Rech.
C. macracanlhum M. Pop.	*C. balcaricum* Galushko	*C. mogollavicum* (M.Pop.jKoroleva	*C. cuneatum* Höchst, ex Rich
C. pungens Boiss.	*C. baldshuanicum* (M.Pop.)Lincz.	*C. monibrelii* Jaub. & Sp.	*C. echinospermum* P.H. Davis
C. rechingeri Podlech	*C. canariense* Santos Guerra & Lewis	*C. multijugum* van der Maesen	*C. Judaicum* Boiss.
C. stapfianum K.H. Rech.	*C. feJlschenkoi* Lincz.	*C. nuristanicum* Kitamura	*C. pinnatifidum* Jaub. & Sp.
C. tragacanthoides Jaub. & Sp.	*C. flexuosum* Lipslcy	*C. oxyodon* Boiss. & H oh.	*C. reliculatum* Ladiz.
	C. floribundum Fenzl.	*C. paucijugum* (M.Pop.)Nevski	*C. yamashitae* Kitamura
	C. graecum Orph.	*C. rassuloviae* Lincz.	
	C. grande (M.Pop.) Korotk.	*C. songaricum* Steph. ex DC.	**Section Chamaecicer**
	C. heterophyllum Contandr. et al.	*C. spiroceras* Jaub. & Sp.	*C. chorassanicum* (Bge) M. Pop.
	C. isaurkum P.H. Davis	*C. subaphyllum* Boiss.	*C. incisum* (Willd.) K.Maly
	C. kermanense Bornm.	*C. laetum* Rassulova & Sharipova	

Section Monocicer includes all the annual species of the genus including the chickpea progenitor and is thus the most important for breeders. There is no controversy about the wild progenitor of chickpea, as all evidence points to *C. reticulatum* as the most likely candidate, based on seed storage proteins (Ladizinsky and Adler 1976; Ahmad and Slinkard 1992), interspecific hybridization (Singh and Ocampo 1993), DNA marker analysis (Penmetsa et al. 2016; Iruela et al. 2002; Sudupak 2004; Sudupak et al. 2002; Patil et al. 1995; Javadi et al. 2007), Karyotype (Ocampo et al. 1992) and isozyme patterns (Labdi et al. 1996).

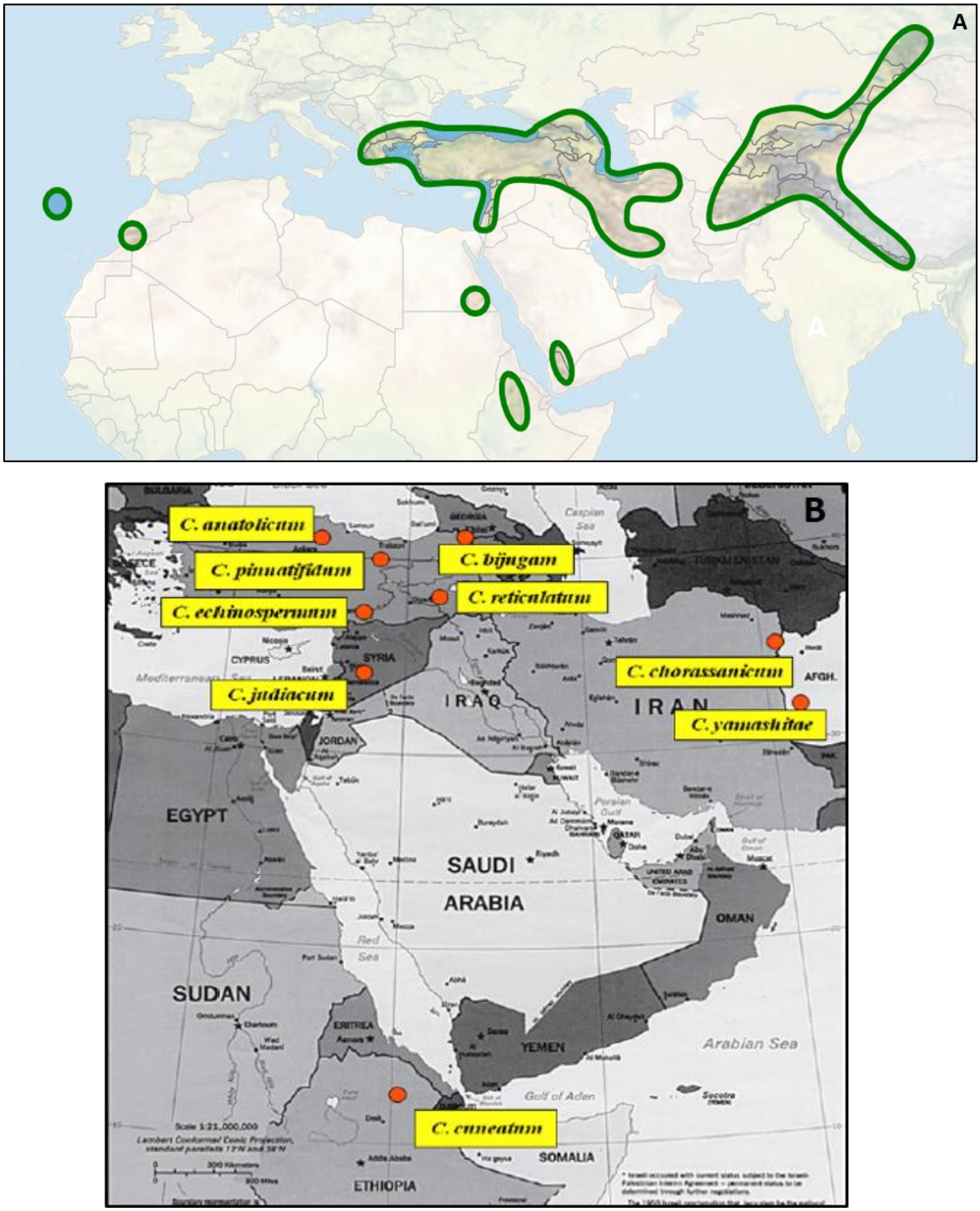


Figure 1.1 (**A**) Geographical distribution of the wild species belonging to *Cicer* genus, according to Van Der Maesen et al. (2007). (**B**) Wild Cicer species found in the chickpea domestication area (Kanouni et al. 2011).

Within chickpea, two well-known market classes can be differentiated according to morphological characters (Fig 1.2). The *kabuli* type is common in the Mediterranean basin and possess white flowers and have large, round or "rams-head" shaped seeds that are pale in colour with a relatively smooth surface. The *desi* type, mainly grown in south-west Asia and

east Africa, accounts for about 70% of all chickpea production, and is characterized by smaller, angular seeds with a harder, darker seed coat, and also smaller leaflets and overall plant stature. They also differ in pigmentation; whereas *kabuli* flowers are white, *desi* flowers are generally purple and diversity for flower, pod, seed and vegetative colour and in seed surface and shape is much wider than in the *kabuli* type (Van Der Maesen 1972). These groups have been traditionally accepted as genetically distinct (Moreno and Cubero 1978), but a recent phylogenetic analysis using a genome-wide marker set was unable to distinguish them, and showed that closely-related *desi* and *kabuli* types were represented in all major germplasm groups (Penmetsa et al. 2016). The white flowers and light-coloured seed coat associated with kabuli types are due to mutations that impair the function of a basic helix–loop–helix (bHLH) transcription factor that acts in part to regulate anthocyanin biosynthesis. The existence of several different mutations in distinct *kabuli* lineages supports the previous idea that *kabuli* types may have originated from *desi* types after domestication (Gil and Cubero 1993; Ladizinsky and Adler 1976), and suggests that this may have occurred several times independently (Penmetsa et al. 2016).

Figure 1.2 *Kabuli* (top) and *desi* (bottom) chickpea flower and seeds. *Desi* flower picture taken from https://daot.tk/bengal-gram-plant, and *kabuli* seed picture taken from https://www.grainews.ca.

1.2 Origin and history.

Evidence of domesticated forms can be found in archaeological sites dated as early as 7500-6800 BC (Zohary and Hopf 2000; Van Der Maesen 1972) This indicates that chickpea was one of the first grain legumes to be domesticated, as one of the eight founder crops that first appeared in the Fertile Crescent region during Neolithic period (Redden and Berger 2007; Abbo et al. 2003b). Within this broader region, chickpea is most likely to have originated in an area corresponding to present-day south-eastern Turkey and adjoining areas of Syria, where its wild progenitor can still be found today (Fig 1-B)(Lev-Yadun et al. 2000; Van Der Maesen 1987). From this point of origin, chickpea is thought to have first spread west to modern Greece in the late Neolithic, and then to the rest of Mediterranean basin in the Bronze age (Fig 1.3). This was a critical period in chickpea history; no archaeological records can be found between 6000 and 4000 BC, in contrast to the continuous records obtained for the other founder crops, suggesting that chickpea cropping suffered a huge decline, which was likely a consequence of Ascochyta blight (Kumar and Abbo 2001; Abbo et al. 2003a). This disease, caused by the fungi *Ascochyta rabiei* (Pass.) Lab., affects the aerial parts of the plant and can be devastating to the point of producing total crop loss in wet environments or winter-grown chickpea (Singh and Reddy 1996; Jayakumar et al. 2005; Daba et al. 2016a). The most accepted theory is that its reappearance as a significant crop in the early Bronze Age was enabled by a shift in the sowing date from autumn to spring, made by the ancient farmers throughout the Mediterranean basin in an attempt to fight this disease (Abbo et al. 2003a). This conversion into summer crop is suggested to have not only provided an escape from Ascochyta blight in those areas with a relatively dry spring but also facilitated its introduction in India around 2000 BC as a post-rainy season crop where it experiences shorter daylengths during the growing period (Allchin 1967; Van Der Maesen 1987). This introduction was highly successful, as shown by the fact that the Indian subcontinent today accounts for 70% of world production and the large variety of landraces, names and recipes found throughout this region.

Ethiopia is considered as a secondary centre of chickpea diversity (Van Der Maesen 1987), and its introduction to this region is thought to have taken place during the Iron Age. The voyage of chickpea through the world had a long pause at this point, until it was introduced in America by the Spanish and Portuguese colonizers in the 16th century (Redden and Berger 2007), and *desi* cultivars to Kenya by Indian immigrants during late 19th century (van der Maesen, 1972).

Due to this atypical history, *Cicer arietinum* is a species with an extremely low genetic diversity. Low genetic diversity is to an extent a common feature of all crops, thought to result from the "founder effect" associated with domestication and the more recent replacement of local landraces by elite cultivars generated with modern breeding techniques. Chickpea, however, suffered two additional bottlenecks during its evolution that highly constrained its genetic basis, even compared with other crops domesticated in the same region during the Neolithic period (Abbo et al. 2003a).

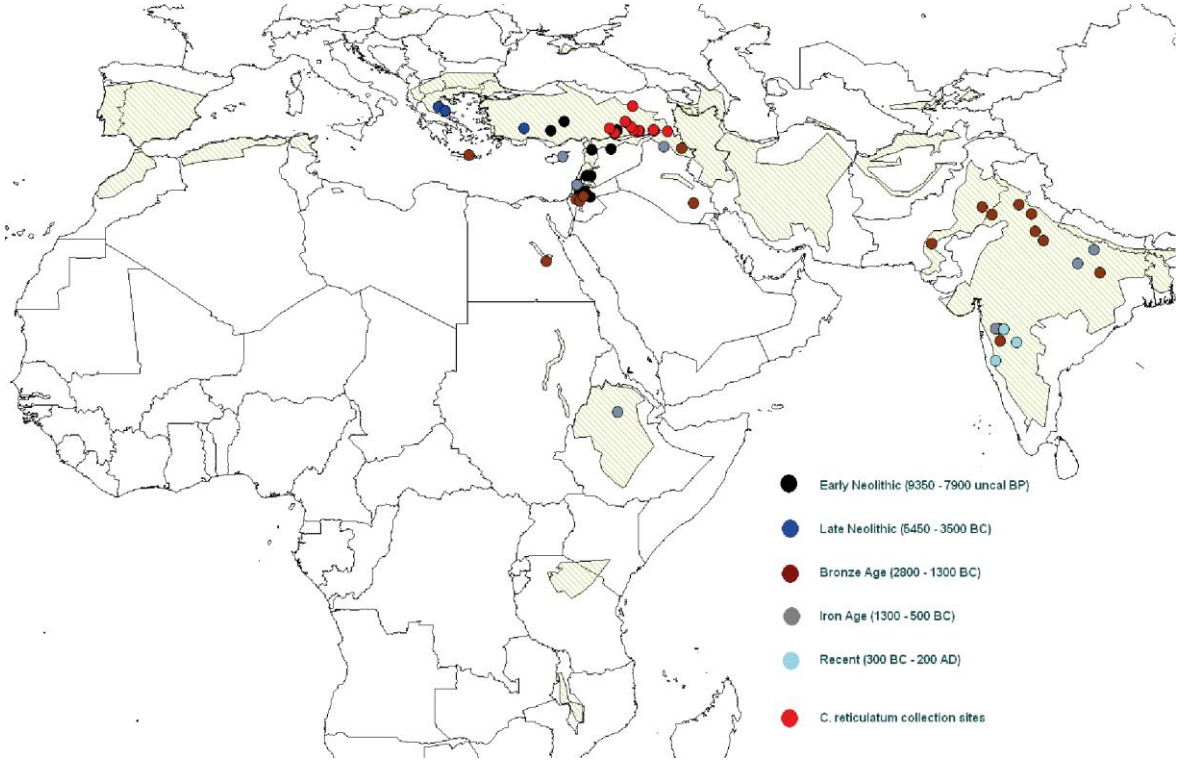


Figure 1.3 Dissemination of chickpea across the old world, from narrowly distributed Mediterranean winter annual (*C. reticulatum*, the wild progenitor) to widespread crop (crosshatched shading), based on archaeological sites containing ancient chickpea or Cicer species (Berger 2014).

Even more recently, the high demand in traditional producer countries like India, Pakistan and Spain has led to the development of an international market and driven the introduction and expansion of chickpea production in countries such as Canada and USA, and also Australia, where the first cultivar was released in 1978 (Croser et al. 2003a; Siddique et al. 2000; Dusunceli et al. 2007).

1.3 Importance of chickpea as a crop.

Predictions based on current growth rates suggest that world population will reach 9.2 billion by 2050, requiring an estimated 70% increase in food production. Legumes in particular are considered important in this future scenario due to their nutritional properties and value as a staple food, especially for subsistence and small-scale farmers in developing countries (Varshney et al. 2013b). Chickpea has several specific advantages that make it particularly valuable in regions where land and water are scarce. From a nutritional point of view, chickpea provides a source of both carbohydrate and high-quality protein, which is particularly valuable for groups with a vegetarian diet, and does not contain the antinutritional factors that are present in many other legumes (Wood and Grusak 2007; Kumar and Abbo 2001). It is also rich in fiber and minerals (phosphorus, calcium, magnesium, iron and zinc) and its lipid fraction is high in unsaturated fatty acids (Jukanti et al. 2012; Mallikarjuna et al. 2011). Chickpea, like other grain legumes, is also used as food for livestock and has a significant role in farming systems as a substitute for fallow in cereal rotations, where it contributes to the sustainability of production and reduces the need for N fertilization through fixing atmospheric nitrogen thanks to its association with bacteria belonging to *Rhizobium* genus (Yadav et al.). Finally, its robust root system allows chickpea to be grown using residual moisture, enabling farmers to practice double cropping, which increases productivity and provides an extra source of income via its trade in both domestic and international markets to satisfy the world growing demand (Kassie et al. 2009).

Nowadays, chickpea is the third most important pulse crop in the world after common bean (*Phaseolus vulgaris* L.) and pea (*Pisum sativum* L.), with a total production of *ca* 13.7 million tonnes from a cultivated area of nearly 14 million ha in 60 countries across the world except in Antarctica (FAOSTAT, 2014). The Indian subcontinent is responsible for the majority of this production (approximately 77%), while the rest is distributed across the Mediterranean basin (including southern Europe), eastern Africa, Australia and both North and South America (Berger and Turner 2007). This wide global spread means that chickpea is subject to very different environmental conditions and cropping systems depending on the economic and agro-climatic conditions of each region where is cultivated; from an industrialized, market-oriented cropping in Canada or Australia to a marginal, post-rainy subsistence farming in arids/semiarid environments. This reflects a wide variation in key environmental variables including photoperiod, temperature and water availability, and the

potential for substantial effects on growth and development and, consequently, on productivity (Kumar et al. 2007).

Globally, chickpea productivity is approximately 0.8 t/ha, far lower than the potential of 3-5 t/ha that can be obtained under optimal circumstances, such as when sown as a winter crop or in favourable conditions with abundant water and fertilizer (Singh et al. 1997; Sadeghipour and Aghaei 2012; Valimohammadi et al. 2007; Iliadis 2001; Siddique and Krishnamurthy 2014), and has increased very little in the last 20 years (Millan et al. 2006; Varshney et al. 2014a). This is due primarily to the constraints imposed by certain biotic and abiotic stresses and also due to the persistent pressure from more productive cereal crops, which since the green revolution in the 1970s have displaced chickpea to marginal, less fertile lands under rainfed conditions. As a result, it is estimated that ninety percent of the world's current chickpea production occurs under conditions which are not optimal to allow it to reach its full potential (Kumar and Abbo 2001). Also, some authors point to the narrow genetic base of chickpea as another factor limiting diversity and adaptive potential in the crop (Millan et al. 2006; Abbo et al. 2003a). For this reason, improvement is very important for the future of the crop, and the main areas of research in chickpea are focused on its major weaknesses. Pests and diseases, especially the fungal diseases Ascochyta blight and Fusarium wilt (caused by *Fusarium oxysporum* f. sp. *ciceri*), have been extensively studied, and high yielding, resistant cultivars have been developed (Gaur et al. 2008; Singh 1997). Abiotic stresses such as terminal drought, heat and salinity are the major abiotic constrains to chickpea productivity, which can result in considerable yield loss and can even cause total crop failure (Kumar and Abbo 2001). Among these, drought is the most commonly-encountered challenge across the different chickpea growing regions globally (Devasirvatham et al. 2012; Zhang et al. 2000; Turner et al. 2001; Berger and Turner 2007).

To overcome these constraints, a number of diverse strategies can potentially be adopted. For example, the introgression of alleles from wild species has been successfully used to improve chickpea resistance to both biotic and abiotic stresses and also to improve yield (Singh et al. 1993; Singh and Ocampo 1997; Singh et al. 2005; Croser et al. 2003a). Another approach is the development of varieties with an early-flowering and maturing phenology that allows the avoidance of unfavourable conditions (Kumar and Abbo 2001; Berger 2007). Flowering time is thus one of the better studied physiological traits in chickpea.

1.4 Flowering in chickpea

1.4.1 Importance of flowering time in chickpea production

Flowering time is crucial in any plant species, as a correct timing of flowering ensures reproductive success, and natural variation in flowering allows species to adapt to different environments (Berger et al. 2006; Berger et al. 2004b). In crops there is an additional economic consideration, since the timing of growth, flowering and maturity have a major impact on final yield. Thus, improving the "fit" of the chickpea life cycle to local growing conditions can lead to higher crop yields.

One of the main difficulties faced by chickpea farmers worldwide is the necessity to complete the crop cycle within a very short season. In semi-arid and tropical environments, chickpea is sown after the rainy season, with only a short growing period before it faces terminal drought. Another drastic example is higher-latitude continental temperate environments like western Canada, where chickpea growth season is only 110-120 days, and the end of the reproductive phase coincides with declining temperatures, resulting in delayed maturity and increased risk of frost damage in the very cold-sensitive phase of pod development (Berger et al. 2004b; Clarke and Siddique 2004; Croser et al. 2003b; Anbessa et al. 2007).

In Mediterranean environments is well-documented that chickpea is most productive when grown as a winter crop, increasing yield between 23 and 188% (Lichtenzveig et al. 2006; Iliadis 2001; Özdemir and Karadavut 2003; Singh et al. 1997). However, spring-sowing is the most common cropping system for chickpea in such environments. This practise helps the crop to avoid fungal diseases but exposes it to hydric stress during the sensitive pod-filing phase, with the consequent risk of terminal drought that can result in yield loss.

In any case, an early phenology have been proposed as the best strategy to adopt in all these short-season environments, since allows the crop to escape terminal stresses and thus evading seed loss (Johansen et al. 1997; Jamalabadi et al. 2013; Kumar and Abbo 2001). This mechanism has been successfully proved as a drought scape in other species such as wheat, barley and maize (González et al. 1999; Mondal et al. 2013; Ngugi et al. 2013). In chickpea, this seems also a successful strategy, as a positive correlation has been found between early flowering and yield (Das et al. 2015b; Hamwieh et al. 2013b; Monpara and Dhameliya 2013; Gaur et al. 2014a), which is especially relevant in dry environments (Devasirvatham et al. 2015; Pushpavalli et al. 2015; Hamwieh et al. 2013b; Rubio et al. 2004). However, it has also

been suggested that a too short crop duration can have a penalty on grain yield, due to an insufficient biomass accumulation (Gaur et al. 2014a).

Phenology is a wide term, and flowering time is only one of its components, as it also includes the time taken to produce pods and for the plant to mature and senesce. However, breeders have traditionally used flowering time as a simple measure of crop duration (Kumar and Abbo 2001), and this is understandable since a positive correlation between early flowering and early maturity have been repeatedly reported (Varshney et al. 2014a; Anbessa et al. 2007; Gaur et al. 2014a), with special significance in short-season environments such as those described above (Subbarao et al. 1995; Kumar and Abbo 2001; Rubio et al. 2004; Anbessa et al. 2007). However, the existence of cultivars with early flowering but late maturity and *vice versa* suggests that although these traits are related, they can be separated to some extent (Summerfield and Roberts 1988), and this offers possibilities for the design of new crop ideotypes.

With this in mind, it is understandable that major chickpea improvement initiatives by breeders and research centres worldwide including the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) and the International Centre for Agricultural Research in the Dry Areas (ICARDA) include a focus on the development of early flowering cultivars, with the aim to develop new cultivars with improved phenology that are better adapted to different environments and hence have a higher yield. However, despite this importance, the molecular understanding of flowering in chickpea is limited relative to other crop species, and there is a strong case for research to better understand the genes and genetic interactions that contribute to variation in phenology.

1.4.2 Physiology of flowering in chickpea

Phenological studies in chickpea reveal that both photoperiod and temperature have an independent but additive effect on flowering time, and models describing flowering time as a function of mean temperature and photoperiod has been proposed by several researchers (Summerfield et al. 1981; Ellis et al. 1994; Roberts et al. 1985; Roberts et al. 1980; Nanda and Chinoy 1960; Daba et al. 2016b; Soltani et al. 2006). This suggest that the responses to these two factors are probably genetically independent, as reported in other species like Arabidopsis (Balasubramanian et al. 2006), lentil (Erskine et al 1990) or rice (Luan et al. 2009).

Regarding photoperiod, chickpea has generally been considered as a quantitative (facultative) long-day plant (Van Der Maesen 1972; Sethi et al. 1981), although more recent studies indicate that it would be more accurate to define it as a qualitative (obligate) long-day plant with a critical photoperiod of 11-12 hours under which it does not flower (Soltani et al. 2004). Daba et al. (2015) proposed 3 categories of photoperiod sensitivity in chickpea by comparing the flowering response of 100 chickpea accession to long (16 hours light) and short days (10 hours light). Day-neutral accessions have less than 10 days difference in flowering date in these contrasting photoperiods, while accessions with a 10-40 day difference were classified as intermediate and those with a difference of >40 days as highly-sensitive. A subsequent study proposed the existence of three flowering induction phases: a photoperiod-insensitive pre-inductive phase, a photoperiod-sensitive inductive phase and a photoperiod-insensitive post inductive phase, with an inverse relationship existing between photoperiod and duration of the sensitive phase (Daba et al. 2016c).

Work on thermal response indicates that progress to flowering is a positive linear function of average temperature between a critical minimum and an upper optimum. Temperatures above this optimum considerably delay flowering, and chickpea is also very sensitive to chilling temperatures at all developmental stages, especially at reproductive stage when temperatures as moderate as 15°C can lead to flower and pod abortion (Clarke and Siddique 2004; Croser et al. 2003b; Ellis et al. 1994; Summerfield and Roberts 1988; Soltani et al. 2006). Another important issue related to temperature is vernalization. Wild species within the genus *Cicer* including *C. reticulatum*, the wild ancestor of *C. arietinum* possess a strong vernalization requirement, (Berger et al. 2004a; Abbo et al. 2002; Berger et al. 2005; Sharma and Upadhyaya 2015) similar to other galegoid legumes Medicago, lentil and pea. However, in domesticated chickpea, the existence of a vernalization response has been widely discussed and is somewhat controversial. Early studies found that exposure to cold early in development tended to promote flowering, and this effect was more evident in short photoperiods (Nanda and Chinoy 1960; Pal and Murty 1941; Angus and Moncur 1980; Saxena and Siddique 1980). However, this was later considered as an artefact that largely reflected a slower growth rate under the cold conditions, and *C. arietinum* was subsequently considered to be insensitive to vernalization (Abbo et al. 2003a; Berger et al. 2012; Summerfield et al. 1989). This idea has been incorporated into an influential hypothesis about evolution of chickpea production, with the loss of vernalization response considered to be a key change enabling a prehistoric shift from winter to summer cropping (Abbo et al. 2003a;

Berger et al. 2012; Summerfield et al. 1989). However, although the vernalization response does appear to have been lost in many early-flowering varieties, other recent reports suggest that vernalization response might still exist in late-flowering germplasm (Sharma and Upadhyaya 2015), and a weak response of two *C. arietinum* cultivars was also observed by (Pinhasi van-Oss et al. 2016). It is clear that further work is required to clarify the existence and extent of vernalisation response in chickpea.

1.4.3 Variation in chickpea flowering time

A substantial range in flowering time, as consequence of different responses to photoperiod and temperature, can be found among cultivated chickpea germplasm (Pundir et al. 1988; Daba et al. 2015; Daba et al. 2016b).

As in the case of other legumes (Erskine et al. 1990), a close association between the geographic origin of cultivars and their flowering time has been reported in chickpea. Cultivars from higher latitudes are mostly late flowering, whereas types originating from lower latitudes generally have an early phenotype. In this respect, Or et al. (1999) found that a collection of Ethiopian chickpea landraces flowered earlier compared to another with Mediterranean origin when grown in Israel. This correlation could be explained by an inverse photothermal sensitivity in the germplasm from different latitudes (Kumar and Abbo 2001; Roberts et al. 1985); accessions with strong photoperiod response, such as those belonging to Mediterranean germplasm, show a weak temperature response. In such environments, photoperiod sensitivity was favoured because temperature raise together with day length, and thus when photoperiod requirements are fulfilled, both demands are satisfied simultaneously. By contrast, in South Asian material where the seasonal day length variation is minimal, more photoperiod-insensitive lines but with a stronger temperature response were selected, which ensures flowering in an optimal temperature. Similar observations were reported by Hovav et al. (2003), and confirmed in more recent studies (Berger et al. 2011; Daba et al. 2015; Pinhasi van-Oss et al. 2016; Daba et al. 2016b).

1.4.4 Genetics of flowering control in chickpea

Despite fundamental interest and applied significance of flowering time in chickpea, relatively little is known about its genetic control. Several studies have addressed this question in classical genetic analyses and QTL studies, and results indicate the involvement of several genes, although the variety of parental lines and marker systems and the strong

influence of environment have made it difficult to compare these studies definitively. Nevertheless, it appears that a small number of major loci may account for most of the phenotypic variation, with further contributions from a larger number of loci with minor effect, similar to the situation in other legumes such as pea and soybean, and more generally in many crops. The studies of Or et al (1999) and Kumar and Van Rheenen (2000) both identified a single major locus, while the results of (Hegde 2010; Anbessa et al. 2006; Ezzat et al. 2015) and (Aryamanesh et al. 2010) all indicate the presence of two loci with additive effects. Gumber and Sarvjeet (1996) also found two loci but with an apparently epistatic interaction, potentially indicating the presence of at least one locus distinct from these other studies. Finally, diallelic analysis using extra-early lines indicated more than two complementing genes operate to control flowering time (Kumar and Abbo 2001).

On the basis of these studies, four major genes have so far been proposed. *Photoperiod* (*Ppd)* is proposed to be a major locus controlling the photoperiod response, with the recessive *ppd* allele (present in line ICC 5810) conferring photoperiod insensitivity (Or et al. 1999). Kumar and Van Rheenen (2000) identified another major locus, *Early flowering 1* (*Efl1*), with the recessive allele *efl1* from accession ICCV2 conferring early flowering. The possibility that these genes might be allelic remained open until complementation experiments of Hegde (2010) revealed they were in fact different genes. This same study described *Efl3*, a new major locus conferring earliness in the cultivar ICC 5810. Finally, Gaur et al. (2014a) described recessive alleles at an additional locus (*Efl4*) conferring earliness in the lines ICC 16641 and ICC 16444.

Although partial dominance of early flowering has been reported in at least one genotype (Monpara and Dhameliya 2013), it seems to be more common in chickpea that alleles conferring earliness are recessive (Anbessa et al. 2006; Aryamanesh et al. 2010; Gumber and Sarvjeet 1996; Hegde 2010; Kumar and Van Rheenen 2000; Or et al. 1999), and therefore it seems likely that variants at all four of these loci represent loss-of function mutations.

Besides genetic studies, considerable effort has been made to map flowering genes through QTL analysis since the development of chickpea genetic maps in the 1990s. The first mapping of flowering time in chickpea was achieved by Cho et al. (2002), who identified a QTL for days to 50% flower in LG3. Since then, diverse linkage analysis have reported more than 50 QTL distributed across all eight chromosomes (summarized in Table 1.2), confirming the existence of several genes involved in the genetic control of flowering in chickpea.

A better genetic understanding of flowering time control is likely to provide significant benefits for chickpea improvement (Kumar and Abbo, 2001). Although not strictly necessary for the development of new early genotypes by traditional breeding, more detailed understanding of the molecular basis for existing variation and the genetic and physiological mechanisms controlling the floral transition will be important to guide the development of new cultivars better adapted to each local environment and cropping system, particularly in response to current challenges such as climate change. Recent studies, such as that of Upadhyaya et al. (2015) are beginning to show how an integrative approach incorporating both genetic and genomic tools may provide an efficient strategy to find candidate genes for major flowering time QTLs. Genomic resources have been developed for chickpea in recent years (reviewed in section 1.6), that will greatly assist the physical positioning of genomic regions of interest, marker design, and the identification and evaluation of candidate genes.

Table 1.2 Summary of QTLs described for flowering time and maturity in chickpea to present day

Cross		Trait	Year	Condition	Place	LOD	PEV (%)	Markers	Reference
Linkage group 1									
Narrow	HadasxICC5810	First flower	2002	Field	Massuot-Yitzchack (Israel)	9.0	56	GAA40, H1F022	(Lichtenzveig et al. 2006)
					Gilat (Israel)	8.8	53		
	ILC588xILC3279	50% flower	2006		Tel Hadya (Syria)	7.8	15	H5A08-TA8	(Rehman et al. 2011)
	ICC283xICC8261				Patencheru (India)	-	8.41	CaM0694 - NCPGR90	(Varshney et al. 2014a)
			2010, 2011				5.84-6.40	NCPGR136 - TR43	
					Nandyal, Hiriyur (India)		6.63-15.57	TA103II - TA122	
	ICC283xICC8261	Maturity			Multilocation (India)		6.08-13.77	TA103II - TA113	
			2011		Nandyal (India)		6.0-7.09	NCPGR136 - TR43	
	ICCV2xJG11		2013-2015	Pots in field	Patancheru (India)	40.76	66.75	CaM1301-CKAM1971	(Pushpavalli et al. 2015)
	ICC5810xCDCFrontier	First flower		Field		12.88	20.28	TA122-TA30	(Mallikarjuna et al. 2017)
Wide	ICC4958xICC17160	50% flower	2012		Multi-location, India	9.8	21.8	CID_C_403402-CID_C_492721	(Das et al. 2015b)
		Maturity				9.5	20.5		
		50% flower	2013			10.7	23.6		
		Maturity				10.1	22.7		
Linkage group 2									
Narrow	HadasxICC5810	First flower	2002	Field	Massuot-Yitzchack (Israel)	4.4	22	H4B09, H1B06	(Lichtenzveig et al. 2006)
					Gilat (Israel)	3.7	18		
	ILC3279xILC588	Flowering	2009, 2010		Sanandaj (Iran)	3.11	23	GAA47-TA37	(Ezzat et al. 2015)
Wide	ICC4958xICC17160	50% flower	2012		Multi-location, India	8.5	17.5	CID_C_4546528-CID_C_4703718	(Das et al. 2015b)
		Maturity				8.1	16.4		
		50% flower	2013			9.7	20.4		
		Maturity				9.4	19.6		
Linkage group 3									
Narrow	ICCV2xJG62	50% flower	1998, 1999	Field	Patancheru (India)	3.03	-	Ts57, TA127	(Cho et al. 2002)
		First flower				2.34			
Wide	ICCL81001xCr59	50% flower	2001	Field	Cordoba (Spain)	5.9	52	TA142	(Cobos et al. 2009)
		Flowering time		Glasshouse[a]		16	26		
Narrow	ICCV96029xCDCFrontier	50% flower	2011, 2012, 2013	Field	Saskatchewan (Canada)	5.3	9	CAV1SC48.1P396061	(Daba et al. 2016a)
		Maturity				5.4	19		
		Flowering node		Cabinet[b]		5.8	11	scaffold1777p70396	
Wide	ICC 3996xILWC 1847	First flower	-	Glasshouse[c]	Western Australia	32.4[d]	90.2[d]	TAA142-TA64	(Aryamanesh et al. 2010)
								TS29-TA76	
Narrow	ICC399x6S95362	Flowering time	2005	Field	Victoria (Australia)	26.1/5.6[e]	21	TS19-TR56	(Hossain et al. 2010)
						29.0/6.2[e]	23		
	ILC588xILC3279	50% flower	2006		Tel Hadya (Syria)	10.9	22	TA6-NCPGR12	(Rehman et al. 2011)
	ICCV2xJG62		2005		Patancheru (India)	2.5	10.1	TA106-Podnode	(Vadez et al. 2012)
			2007			3.5	13.6	TA14s-TR40	
	ILC588xILC3279	50% flower	2008-20111		Terbol (Lebanon)	6.07-8.64	17.7-24.2	H6C-07	(Hamwieh et al. 2013b)
					Tel Hadya (Syria)	2.04- 3.94	5.1- 9.7	H1F-14; H4G-07	

Table 1.2 Continued

Cross		Trait	Year	Condition	Place	LOD	PEV (%)	Markers	Reference
Linkage group 3 (Continued)									
Narrow	ICC4958xICC 1882	50% flower	2008	Field	Patancheru (India)	-	9.22	TR2 - H3C06	(Varshney et al. 2014a)
	ICC283xICC8261		2005, 2006			-	13.78-18.97	CaM1753 - cpPb-677529	
	ICC4958xICC 1882		2009		Durgapura (India)	-	8.08-8.23	TA76s - TR24	
		Maturity					8.58		
	ILC3279xILC588	Flowering	2009, 2010		Sanandaj (Iran)	3.49	35	STMS25-TR59	(Ezzat et al. 2015)
Wide	ICC4958xPI489777	Vernalization			Patancheru (India)	27	47.9-54.9	CaM0717-TA64	(Samineni et al. 2016)
	ICC4958xICC17160	50% flower	2012		Multi-location, India	10.3	24.7	CID_C_4424175-CID_C_4575860	(Das et al. 2015b)
		Maturity				9.8	23.5		
		50% flower	2013			11.5	27.5		
		Maturity				11.2	26.9		
Narrow	ICC16374 x ICC762	50% flower	2012	Field	Patancheru (India)	5.4-6.3	11.6-12.3	CaKSNP4307-CaKSNP4313	(Upadhyaya et al. 2015)
			2013	Glasshouse	New Delhi (India)	7.5-8.9	14.3-15.4	CaKSNP4801-CaKSNP4805	
	ICCV96029xCDCFrontier	First flower	2013-2015	Field	Patancheru (India)	3.45	5.66	CaM1122-TR13	(Mallikarjuna et al. 2017)
	ICC5810xCDCFrontier					16.7	24.95	CaM1358-TA142	
	BGD132xCDCFrontier					5.24	4.39	CaM1515-TR13	
						4.21	4.04	TA142-TA64	
Linkage group 4									
Narrow	CA2156xJG62	50% flower	2003	Field	Cordoba (Spain)	2.4	12.4	GAA47	(Cobos et al. 2007)
			2004			4.4	20.3		
				Glasshouse[a]		4	23		
	ICCV96029xCDCFrontier		2011, 2012, 2013	Field	Saskatchewan (Canada)	3.1	11	scaffold 2005p25023	(Daba et al. 2016a)
						5.7	10	scaffold360p479554	
						3.8	13	CAV1SC2.1P566504	
						5.7	14	scaffold34p1977386	
		First flower		Cabinet[b]		3.1	11	scaffold 2005p25023	
		Flowering node				4.2	10	scaffold360p644415	
						9.4	29	CAV1SC2.1P3082421	
	ILC588xILC3279	50% flower	2006	Field	Tel Hadya (Syria)	2.9	5	TA132-GA137	(Rehman et al. 2011)
	ICCV2xJG62		2005	Field	Patancheru (India)	2.6-5.8	13.2-37.7	TA144-NO_Y_13	(Vadez et al. 2012)
						6.2	15.8	TA127-TS57	
			2007			3.3	18.5	TA186-TA45	
	ILC588xILC3279		2008-20111	Field	Terbol (Lebanon)	2.53	7.8	H1B-17	(Hamwieh et al. 2013b)
					Tel Hadya (Syria)	2.15	5.3	H5G-01	
	ICC4958xICC1882		2008		Sehore (India)	-	24.49	NCPGR127 - TAA170	(Varshney et al. 2014a)
		Maturity	2009				19.71		
	ICCV2xJG11		2010	Pots in field	Patancheru (India)	27.99	62.67	ICCM0293-CKAM0707	(Pushpavalli et al. 2015)
		Flowering	2011			2.91-5.73	22.6-38.9	CKAM0003-CKAM1003	
		Maturity				30.34-39.19	59.95-64.34		
	JG62 x Vijay	Maturity	2003-2006	Field, winter	Rahuri (India)	4.3-4.8	14.1-18.7	NCPGR80	(Gowda et al. 2011)

Table 1.2 Continued

Cross		Trait	Year	Condition	Place	LOD	PEV (%)	Markers	Reference
Linkage group 4 (Continued)									
Narrow	ICC16374 x ICC762	50% flower	2012, 2013	Field Glasshouse	Patancheru (India) New Delhi (India)	10.5-11.7 10.8-11.5	24.1-27.3 22.4-24.5	CaKSNP6695-CaKSNP6704 CaKSNP5894-CaKSNP5910	(Upadhyaya et al. 2015)
	ICCV96029xCDCFrontier	First flower	2013-2015	Field	Patancheru (India)	5.66	11.75	GAA47-ICCM192a	(Mallikarjuna et al. 2017)
	ICC5810xCDCFrontier					9.18	10.53	NCPGR21-GAA47	
Wide	ICC4958xPI489777	Vernalization	2009	Field	Patancheru (India)	3	13	ICCeM005-ICCM0065b	(Samineni et al. 2016)
Linkage group 5									
Narrow	ICCV96029xCDCFrontier	50% flower	2011, 2012, 2013	Field	Saskatchewan (Canada)	18	44	CAV1SC1.1p4940145	(Daba et al. 2016a)
	ICCV2xJG62		2005		Patancheru (India)	3.4	13.8	TA114-TA78	(Vadez et al. 2012)
			2007			3.6	12.6		
	ICC283xICC8261	Maturity	2005		Patancheru (India)	-	6.32	CaM1372 - CaM1977	(Varshney et al. 2014a)
			2011		Sehore (India)		16.79	CaM2029 - TA11	
	ICCV2xJG11		2010, 2011	Pots in field	Patancheru (India)	16.16-28.95	40.69-61.06	CaM0463-ICCM272	(Pushpavalli et al. 2015)
		Flowering	2011			6.26-8.01	24.98-26.93		
	ICC16374 x ICC762	Days to 50% flower	2012, 2013	Field, Greenhouse	New Delhi, Patancheru (India)	8.4-9.5	17.9-19.5	CaKSNP8449-CaKSNP10106	(Upadhyaya et al. 2015)
	ILC2379xICCV2	Flowering	2009		Sanandaj (Iran)	5.61	33	TA117-STMS22	(Jamalabadi et al. 2013)
Linkage group 6									
Narrow	ICC4958xICC1882	50% flower	2009	Field	Sehore (India)	-	6.77	CaM1125 - H4H06	(Varshney et al. 2014a)
	ICC283xICC8261		2005		Patencheru (India)		11.85	CaM1257 - ICCeM15	
			2010, 2011		Nandyal, Sehore (India)		5.76-7.70	NCPGR200 - CaM0317	
	ICC4958xICC1882	Maturity	2009		Nandyal, Patancheru (India)		5.94-12.13	TA106 - CaM0399	
	ICC283xICC8261		2005, 2010		Patancheru, Nandyal (India)		2.95-10.47	CaM1257 - NCPGR200	
			2011		Patancheru (India)		8.86	NCPGR200 - ICCeM15	
	Vijay x ICC4958		2002-2006	Field, winter	Rahuri (India)	3.9	10.2	UBC284	(Gowda et al. 2011)
	ICC16641xCDCFrontier	First flower	2013-2015	Field	Patancheru (India)	55.6	88.19	TA14-TR44	(Mallikarjuna et al. 2017)
Linkage group 7									
Narrow	ICCV2xJG11	Flowering	2010	Pots in field	Patancheru (India)	2.96	6.47	CaM2031-CKAM0165	(Pushpavalli et al. 2015)
		Maturity				4.01	7.95		
	JG62xVijay		2003-2006	Field, winter	Rahuri (India)	3.9	14.4	TA117	(Gowda et al. 2011)
						2.9-3.5	10.2-15.2	TA180	
Linkage group 8									
Narrow	ICC4958xICC 882	50% flower	2005, 2008, 2009	Field	Multilocation (India)	-	9.11-26.87	NCPGR164 - CaM1918	(Varshney et al. 2014a)
	ICC283xICC8261		2005		Patencheru (India)		12.6807	CaM0772 - TS45	
		Maturity	2006, 2008, 2009		Multilocation (India)		6.05-18.83	NCPGR164 - CaM1918	
			2006		Patancheru (India)		8.89	CaM0772 - TS45	
			2005				16.20	GA6 - TA118	
	ICCV2xJG11		2010	Pots in field		29.8	66.97	CaM0812-CKAM0647	(Pushpavalli et al. 2015)
						3.6	13.24	CKAM1903-CKAM0343	
			2011			31.87-43.32	56.87-65.07		
		Flowering				7.75-10.62	37.75-39.97		
	ILC588xILC3279	50% flower	2006	Field	Tel Hadya (Syria)	3.7	8	TA159-GA6	(Rehman et al. 2011)

Table 1.2 Continued

Cross		Trait	Year	Condition	Place	LOD	PEV (%)	Markers	Reference
Linkage group 8 (Continued)									
Narrow	ICCV96029xCDC Frontier	50% flower	2011 2012, 2013	Field	Saskatchewan (Canada)	4.3	17	scaffold937p67148	(Daba et al. 2016a)
		Maturity				3.3	15	scaffold1567p981540	
		Flowering node		Cabinet[b]		9.3	32	scaffold1439p220499	
	ICC5810xCDCFrontier	First flower	2013-2015	Field	Patencheru (India)	17.79	25.73	GA6-TA118	(Mallikarjuna et al. 2017)
	BGD132xCDCFrontier					44.38	64.95	TA127-H1D24	
Wide	ICC4958xPI489777	Vernalization	2009	Field		3	8.7	CaM0814-CAM493	(Samineni et al. 2016)
Unknown linkage group									
Narrow	JG62xVijay	Maturity	2003-2006	Field, winter	Rahuri (India)	2.0-3.5	7.3-18.0	UBC299x	(Gowda et al. 2011)
						3.7	16.3	UBC90y	
						3.5-4.7	11.4-23.2	STMS13	
	Vijay x ICC4958		2002-2006			4.6	17.5	TS54	
						5.1	12.9	UBC346	
						3.7	13.9	NCPGR80	
						3.2	9.1	TA42	
						3	14.8	UBC881	
						2.2-3.1	6.0-9.1	TS12	
						3.2	8.1	UBC77y	

a. Short days. Natural daylight in late November when day duration is around 10 h
b. Growth chamber in Long days photoperiod of 16 hours light, 8 hours dark
c. Natural daylight of approximately 12 hours
d. QTLs with epistatic effect. Combined LOD and PEV are given
e. Likelihood Ratio Statistics (LRS)

1.5 Genetic control of flowering time

The life cycle of flowering plants can be considered as a succession of distinct growth phases, and the transition between these phases is dependent on developmental genetic programs that are triggered and modulated by both environmental and endogenous stimuli (Huijser and Schmid 2011). The transition from vegetative to reproductive phase is, perhaps, the most important developmental switch in any plant species, and its timing is especially relevant in the case of annual plants, where alignment of reproductive development with environmental factors is critical to maximise the chance of reproduction and being represented in the next generation (Levy and Dean 1998). This alignment depends on the detection of external clues such as photoperiod and temperature, and the integration of this information with endogenous factors (e.g. age, energy balance, hormones, and the circadian clock) to ensure that onset of flowering takes place in optimal environmental and physiological conditions.

Understanding of the genetic and molecular control of flowering is most advanced in model species, particularly *Arabidopsis thaliana*. Fortunately, Arabidopsis and legumes are relatively closely related to legumes taxonomically, as both fall within the rosid clade of plants. Thus, in an attempt to understand flowering time in chickpea, it is reasonable to begin with the foundation provided by Arabidopsis. This section will therefore review the current knowledge about key genes involved in the control of floral induction in Arabidopsis, as well as the extent to which these gene functions and interactions may be conserved in legume species.

1.5.1 Arabidopsis

Decades of work in the long day-plant, vernalization-responsive model species *Arabidopsis thaliana* has shown that environmental cues controlling flowering are signalled through multiple genetic pathways and integrated with endogenous signals by a few genes normally referred to as floral integrator genes. The interaction between these integrators will determine the expression level of a mobile signal (florigen) that will lead to the activation of floral identity genes and initiation of floral development (Lifschitz 2014). Multiple lines of evidence all indicate that the identity of this mobile signal in *A. thaliana* is the FT protein, a member of a small family of plant proteins related to phosphatidyl ethanolamine binding protein

(PEBP) signalling proteins in animals (Kobayashi et al. 1999). *FT* expression is induced in the vascular tissue, from where the FT protein migrates to the shoot apex. Within the shoot apical meristem, FT binds the bZIP transcription factor FD to form the 'Florigen Activation Complex' (FAC), which may also include a 14-3-3 protein (Taoka et al. 2011; Corbesier et al. 2007; Abe et al. 2005). Once active, this complex activates floral meristem identity genes and other floral promoters such as *APETALA1* (*AP1*), *FRUITFULL* (*FUL*) and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), which specify the differentiation of cells in the flank of the apical meristem into floral meristems (Wigge et al. 2005). A close *FT* paralog *TWIN SISTER OF FT* (*TSF*) acts redundantly to promote flowering using the same molecular mechanism (Yamaguchi et al. 2005).

Hundreds of genes interact in a very complex network to control *FT* expression in Arabidopsis (Andrés and Coupland 2012; Song et al. 2015), and are traditionally grouped in different pathways. The most important of these are the photoperiod and the vernalization pathways. The molecular signalling in response these two environmental cues converges in each case on an integrator gene that functions upstream of *FT* to directly modulate its expression according to the environment. In the case of the response to photoperiod, this gene is *CONSTANS* (*CO*), a zinc-finger transcription factor that serves to integrate information from photoreceptors and the circadian clock for measurement of photoperiod (Greenham and McClung 2015; Huang and Nusinow 2016; Nakamichi 2011). The Arabidopsis clock is currently understood to involve the co-action of over 25 genes that form several interconnected transcriptional feedback loops and additional associated genes. The clock helps generate a diurnal rhythm in *CO* transcript levels such that *CO* peak expression occurs late in the day in long days but after dusk in short days (Suárez-López et al. 2001). CO protein levels are also adjusted according to information about light quality and intensity perceived through different photoreceptors, including phytochromes and cryptochromes (Somers et al. 1998); CO protein is stabilized by light but degraded in darkness, so the interactions between light and the clock ensure that CO protein only accumulates and CO-dependent *FT* induction occurs only under LD (Valverde et al. 2004).

Temperature is another key environmental variable limiting the distribution of plant species (Wigge 2013). Studies on how temperature affects flowering time have been mostly focused on the vernalization response, which can be defined as the promotion of the competence for

flowering by long periods of low temperatures (Kim et al. 2009). In Arabidopsis, this response is mostly mediated through the previously mentioned MADS-box *FLOWERING LOCUS C* (*FLC*) (Michaels and Amasino 1999, 2001), a floral repressor that delays flowering by preventing the expression of floral integrators such as *FT* and *SOC1* (Searle et al. 2006). Multiple pathways regulate *FLC* expression; *FRIGIDA* (*FRI*) and *FRIGIDA-LIKE* genes are the main transcriptional activators, responsible for establishing the high *FLC* expression that maintains the vegetative state of the plant during winter (Johanson et al. 2000; Michaels et al. 2004; Schlappi 2006). Vernalization promotes flowering through the stable epigenetic silencing of *FLC*, which is achieved through a complex mechanism involving a group of antisense transcripts produced at the *FLC* locus (collectively named *COOLAIR*) and the combined action of many other genes, among which those belonging to the Polycomb repressive complex 2 (*PRC2*; *FIE*, *VRN2*, *MSI1* and *SWN* or *CLF)*, and the plant homeodomain (*PHD*) family *VRN5*, *VIN3* and *VEL1* are especially relevant, since they form a complex responsible for placing the histone marks repressing *FLC* and also for the maintenance of this epigenetic inactive state (Berry and Dean 2015; Berry et al. 2015; Sung and Amasino 2006; Sung and Amasino 2004; Sung et al. 2006). Little is known about the cold sensing mechanism, but in Arabidopsis, the cellular signalling initiated by cold triggers the induction of *VIN3*. This induction is correlated with the duration of cold exposure and the strength of the vernalization response, so it is likely that this gene is essential in the molecular mechanism of vernalization in this species (Sung and Amasino 2004).

Several genes included in the so called "autonomous pathway" also participate in the Arabidopsis vernalization response, but in addition they regulate *FLC* in a parallel and independent manner, through different processes such as RNA processing (*FCA*, *FY*, *FPA* and *FLOWERING LOCUS K*) and chromatin remodelling by histone deacetylation (*FVE* and *FLOWERING LOCUS D*) (Feng and Michaels 2011; Schmitz and Amasino 2007; He 2012; Swiezewski et al. 2009).

In contrast to the detailed understanding of the vernalization pathway, less is known about the molecular basis by which ambient temperature influences flowering time. Nevertheless, a number of genes involved in temperature sensing and response have been identified and their mechanism of action uncovered. One example is the well-studied interaction between *FLOWERING LOCUS M* (*FLM*) and another MADS-domain gene *SHORT VEGETATIVE*

PHASE (*SVP*). Two major *FLM* isoforms exist whose relative abundance depends on ambient temperature. FLM-β is preferentially produced at low ambient temperature, forms a complex with *SVP* that down-regulates the expression of important flowering time genes such as *SOC1* and *FT*. *By contrast*, the FLM-δ isoform is more abundant at higher temperatures and forms a complex with SVP that is unable to bind DNA, thus releasing the flowering repression exerted by SVP/FLM-β (Jeong et al. 2007; Lee et al. 2013; Posé et al. 2013). Genes initially located in other pathways also may have a role in the molecular signalling of temperature. For instance, *fca*, *fve*, *fy* and the *tfl1*/*elf3* double mutant are insensitive to temperature changes (Strasser et al. 2009; Blázquez et al. 2003). Recent findings suggest that *CO* might also be implicated in the *FT* up-regulation in response to temperature (Fernandez et al. 2016). Finally, chromatin composition might be also important for temperature perception. An alternative histone (H2A.Z) is evicted from gene regulatory regions at high temperatures, allowing the binding of transcriptional regulators. The Arabidopsis flowering promotion by high temperatures uses this mechanism to allow the binding of the transcription factor *PHYTOCHROME INTERACTING FACTOR* to the *FT* promoter to up-regulate its expression (Proveniers and van Zanten 2013; Kumar and Wigge 2010; Kumar et al. 2012).

In parallel with external signals, certain metabolic and hormonal signals also have an important influence on regulation of flowering in Arabidopsis (Galvão and Schmid 2014). Several hormones influence flowering by interacting with known floral pathways, and among these, gibberellic acid (GA) is perhaps the most dominant (Conti 2017; Campos-Rivero et al. 2017). The control of flowering by GA signalling is thought to involve multiple mechanisms that modulate the expression of key floral genes such as *FT* (Hisamatsu and King 2008; Porri et al. 2012), *LFY* (Gocal et al. 2001) and *SOC1* (Moon et al. 2003a).

Also especially relevant is the age-dependent pathway, present throughout the plant kingdom (Bratzel and Turck 2015). In Arabidopsis, two major microRNAs (miRNAs) families, play a key role in this process; levels of *miR156* decline as plants become older, while *miR172* levels follow the opposite pattern. They also have antagonist roles in flowering; *miR172* induces *FT* expression by targeting floral repressors like *APETALA2* (*AP2*) and *SCHLAFMÜTZE* (*SMZ*) (Aukerman and Sakai 2003; Jung et al. 2007; Mathieu et al. 2009), while *miR156* prevent precocious flowering by downregulation of *miR172* and many members of the *SQUAMOSA*

PROMOTER-BINDING PROTEIN-LIKE (*SPL*) family (Wu and Poethig 2006). The *SPL* genes are essential to regulate the juvenile to adult phase transition and influence flowering through direct activation of the expression of *FT* and also of meristem identity genes such as *LEAFY* (*LFY*), *FUL* and *AP1* (Blázquez et al. 2003).

1.5.2 Legumes

Most well-known legume species fall into two main clades that differ in how flowering time is regulated. Members of the galegoid clade (such as pea, lentil, and chickpea) originate mainly from temperate regions and respond to vernalization, while members of the phaseoloid clade are typically short-day plants from warmer, lower latitudes (e.g. soybean and common bean) and do not have a requirement for vernalization. Knowledge about flowering time regulation in these two groups is most advanced in pea and soybean, respectively, although the two temperate species *Medicago truncatula* and *Lotus japonicus* have complemented this work.

Recent years have witnessed enormous advances in genomics tools and resources in legumes, and genome sequences are now available for many of them (Pandey et al. 2016). Besides being a platform for the application of many other genomic tools, genome sequences enable a comprehensive analysis of the presence and evolutionary relatedness of relevant genes described in other species (Distelfeld et al. 2009; Andrés and Coupland 2012; Greenup et al. 2009; Shrestha et al. 2014). These analyses have revealed that most genes and gene families described in Arabidopsis are conserved to some extent in legumes. However, there are many cases where duplication, loss or functional change has occurred as a consequence of genome evolution after the divergence of legume and Arabidopsis lineages (Liew et al. 2014b; Jung et al. 2012; Kim et al. 2013b; Kim et al. 2012; Young and Bharti 2012; Weller and Ortega 2015). The lack of the *FLC* clade is perhaps the most striking difference, in view of the central role that these genes play in Arabidopsis (Ruelens et al. 2013; Hecht et al. 2005). Recent studies in Medicago and narrow-leafed lupin suggest that as in Arabidopsis, legume *FT* genes are the ultimate targets of vernalization regulation, but more work needs to be done to discern whether cold treatment acts directly on these genes and/or through intermediates (Nelson et al. 2017; Laurie et al. 2011), and the molecular mechanism for vernalization response in this family still remains largely unknown.

Photoperiod and circadian clock are strongly interconnected, and in Arabidopsis they converge to regulate expression of *CONSTANS* (*CO*). *CO*-like homologs are found in legumes (Liew et al. 2014b), but growing evidence suggests that they may not have a central, conserved role in photoperiod sensing. In Medicago, analysis of eight of the eleven *COL* genes, including the closest *CO* homolog *COLa* showed that none could complement Arabidopsis *co* mutants, and their expression patterns suggests no involvement in floral induction (Wong et al. 2014). By contrast, two soybean *GmCOLs* were able to rescue the late flowering phenotype of Arabidopsis *co-1* mutants, and their overexpression resulted in flowering time alterations (Cao et al. 2015a; Wu et al. 2014), indicating that the function of legume *COL* genes may differ between species.

Despite this major difference in the apparent absence of CO function, evidence indicates that other aspects of the photoperiod mechanism and the circadian clock are nevertheless conserved in legumes and have a similar role in flowering time control. For example, the phenotype of *PHYTOCHROME A* (*PHYA*) mutants in pea (Weller et al. 2004; Weller et al. 1997a) and the identification of a *CYCLING DOF FACTOR* (*CDF*) as *LATE BLOOMER2* by Ridge et al. (2016) suggest a common light perception and signalling mechanism. *LATE1* is the pea ortholog of the important gene *GIGANTEA* (Hecht et al. 2007b), and mutants for the genes forming the evening complex have also been found in pea; *ELF3*, *ELF3b*, *ELF4*, *LUX* are the genes behind the loci *HIGH RESPONSE*, *PHOTOPERIOD*, *DIE NEUTRALIS*, *STERILE NODES* (Liew et al. 2009; Weller et al. 2012; Liew et al. 2014a; Rubenach et al. 2017). Analysis of the functions and interaction of these mutants suggest a circadian clock mechanism in legume species that is overall well conserved compared to Arabidopsis.

In soybean, ten loci control most of the flowering variation (Weller and Ortega 2015). Six of them appear to be specifically involved in photoperiod responsiveness. Of these, the *E3* and *E4* loci are two different *PHYTOCHROME A* homologs that act together to delay flowering under non-inductive LD (Watanabe et al. 2009), and *E2* is a GIGANTEA ortholog, functioning in a similar way to pea *LATE1* (Watanabe et al. 2011). The phenotype of all these soybean mutants suggest that the role of the genes from in different flowering pathways is overall well conserved between galegoid and phaseoloid legumes, always considering the opposite direction that flowering regulation must take in some pathways, derived from the different photoperiod requirement (LD vs SD plants) of both clades (Liew et al. 2014b; Weller and Ortega 2015).

Among endogenous signals affecting flowering, some discrepancies were found; For example, gibberellin-deficient mutants in Arabidopsis are late in long days and unable to flower in short days, indicating an absolute hormonal requirement when the photoperiod pathway is not active (Wilson et al. 1992). By contrast, this requirement is not present in legume flowering, which seems to be unaffected in GA-deficient mutants (Weller et al. 1997b). In the other hand, the aging-pathway in legumes could be conserved in great extent, as proven by recent studies: Overexpression of *miR156* delays flowering in alfalfa and soybean, and in the latter species was shown to also down-regulate *miR172*, consistent with Arabidopsis phenotype (Aung et al. 2015; Cao et al. 2015b).

FT-like genes are not only important in Arabidopsis. They have been found in a wide range of plants, with a floral promotion activity conserved despite varied flowering requirements of distant clades, suggesting a highly conserved floral inducing signal in plant kingdom (Chaurasia et al. 2017; Putterill and Varkonyi-Gasic 2016). As in many species, the PEBP family in legumes has undergone expansion relative to Arabidopsis, and all species examined to date possess multiple *TFL1* and *FT* homologs (Jung et al. 2012; Laurie et al. 2011; Hecht et al. 2011; Ridge et al. 2017). Within the *FT* clade, legumes possess three distinct subclades (*FTa*, *FTb* and *FTc*), that seem to have generally-conserved roles as integrators of environmental cues to trigger flowering although there is some evidence for subfunctionalization and the acquisition of cross-regulation during evolution (Ono et al. 2010; Kong et al. 2010; Laurie et al. 2011; Hecht et al. 2011; Yamashino et al. 2013; Pin and Nilsson 2012; Zhao et al. 2016). These reports also suggest conservation of the floral repressive role in some *TFL1* homologs, while others are associated with meristem determinacy in species of both galegoid and phaseoloid clades (Foucher et al. 2003; Kwak et al. 2008; Liu et al. 2010).

The *FT* mechanism of action in legumes seems to be similar to that described in Arabidopsis; the pea mutant *veg2* was characterized as an *FD* ortholog, and confirmed the interaction of VEG2 with FT proteins to regulate a set of downstream of MADS genes essential for correct inflorescence and flower development, similar to Arabidopsis (Sussmilch et al. 2015). In legumes, mutations in some of these genes revealed conservation in this function; for example, the pea mutant *vegetative1* (*veg1*), unable to flower in any condition, is a divergent homolog of Arabidopsis *FUL* (Berbel et al. 2012). Soybean *Dt2*, a gain-of function allele that causes a

determinate phenotype, has recently been identified as an ortholog of pea *VEG1* (Ping et al. 2014).

In conclusion, despite some differences, the flowering pathways described in Arabidopsis seem to be largely conserved in legumes, which suggests that the use of both Arabidopsis and legume systems may provide a good reference for comparative analysis of flowering control in legume species such as chickpea where no previous molecular information is available. The value of this comparative approach is illustrated by the recent study of Ridge et al (2017) which is the first to identify a chickpea flowering locus to the molecular level, presenting evidence that *Efl1* is most likely the ortholog of Arabidopsis *ELF3*.

1.6. Genomic resources in chickpea

Despite its globally high production, nutritional value and an increasing economic importance (see section 1.3), chickpea has suffered from a lack of genomic resources that has limited breeding to conventional methods. Fortunately, enormous progress has been made in recent years that has opened up a new range of genetic resources and approaches that promise to drive rapid advances in understanding of chickpea biology and accelerate the development of new cultivars (Gaur et al. 2014b).

The first genetic map constructed in chickpea using molecular markers included RFLP and RAPD markers along with isozyme markers (Simon and Muehlbauer 1997). Since then, numerous linkage maps have been developed involving several crosses and more reliable markers such as simple sequence repeat (SSR) and sequence tagged microsatellite markers (STMS). These have opened up the possibility to unify linkage group nomenclature and compare maps across populations and have provided anchor points to study the synteny with other legume species (Millan et al. 2014). Over time, the application of sequencing methods and resources has facilitated the development of new types of markers like expressed sequence tags (EST), cleaved amplified polymorphic sequences (CAPS), Diversity Arrays Technology (DArT) or single nucleotide polymorphism (SNP), and this has enabled the construction of improved genetic and physical maps (Varshney et al. 2013a; Varshney et al. 2009c; Upadhyaya et al. 2011; Zhang et al. 2010). These markers and maps have also been utilized for association studies linking interesting traits to specific parts of the chickpea genome, and this has resulting in a large number of

Quantitative Trait Loci (QTL) studies addressing traits such as biotic and abiotic tolerance, phenology and yield improvement (Millan et al. 2006; Gaur et al. 2012).

Transcriptome sequencing has been another important advance in the chickpea genomic race, with reference transcriptomes developed not only for the cultivated species but also for the wild *C. reticulatum* (Hiremath et al. 2011; Varshney et al. 2009b; Garg et al. 2011; Jhanwar et al. 2012; Singh et al. 2013). However, the ultimate molecular resource for any species is the full sequence of its genome, as it dramatically facilitates the identification of genes and functional elements and provides the genomic tools and platforms for gene mapping, isolation and molecular breeding (Varshney et al. 2013a). In the case of chickpea this was achieved in 2013 for both market classes with the *kabuli* variety CDC Frontier and the *desi* landrace ICC4958 (Varshney et al. 2013c; Jain et al. 2013). It has subsequently enabled the resequencing of other lines and the use of genome-wide association studies (Thudi et al. 2016; Li et al. 2017; Das et al. 2015a; Upadhyaya et al. 2015; Upadhyaya et al. 2016) and has opened up the possibility to apply a range of genomic-assisted breeding approaches (Varshney et al. 2013a).

1.7 Aims and scope of this study

With the recent dramatic expansion of genomic tools and resources in chickpea, there rapidly increasing interest in genetic and genomic analysis of key production traits. This thesis investigates flowering time in chickpea, and aims to provide new understanding of how it is controlled at the genetic, molecular and physiological levels.

In Chapter 3, the newly-available genome sequence was used to investigate the conservation of important flowering genes in chickpea and characterize the major flowering gene families. This information was integrated with that from published genetic studies to examine the colocation of flowering genes and QTL controlling flowering and identify potential candidate genes.

Chapter 4 presents the genetic analysis of a major flowering time locus on chickpea chromosome 3 that specifies a major difference in flowering behaviour of wild and cultivated chickpea. It also describes the evaluation of several genes in the QTL region by mapping and expression analysis.

Late flowering in wild *C. reticulatum* is associated with a prostrate growth habit and a profusion of branches. Chapter 5 will investigate the genetics of branching and growth habit in the same material used in chapter 4, and assess the relationship with the major flowering time locus.

Chapter 6 will explore the sequence variation within a cluster of candidate genes underlying the major flowering QTL in chromosome 3, in a chickpea collection of diverse flowering phenotype and geographical origin, in an attempt to identify a possible molecular basis for the QTL.

Chapter 7 addresses the somewhat controversial topic of vernalization responsiveness in chickpea, examining the vernalization response of several cultivated and wild chickpea accessions under controlled environmental conditions. It also investigates the molecular basis of the vernalization response in legumes and the possible role of the chromosome 3 QTL.

This thesis concludes with a general discussion (Chapter 8), in which the major findings from each individual chapter will be evaluated. Their overall contribution towards the global understanding of the control of flowering in chickpea will be summarized, some potential implications for agronomic practices will be indicated and several potential future lines of research will be outlined.

Chapter 2. General Materials and Methods

This chapter describes the materials and methods used for all research presented in this thesis. Individual chapters include a material and methods section with details for specific experiments.

2.1 Plant growth conditions.

All seeds used in this thesis were scarified and coated with a fungicide (Thiram) prior to sowing. Seed were sown in 14 cm pots containing a 1:1 mixture of granulated sand and pasteurised potting mix topped with a 3 cm potting mix layer. Plants were water regularly and received nutrient solution weekly.

For all the experiments described in this thesis, plants were grown in phytotrons at the University of Tasmania, with a diurnal temperature of approximately 24^{o}C and 16^{o}C during the night. Unless otherwise indicated in any individual chapters, photoperiod conditions for the different experiments were as follows: Plants under short day (SD) conditions received 8 hours of natural daylight and 16 hours of total darkness, while those in long day (LD) received natural daylight extended before dawn and after dusk with artificial light (50µmolm-2s-1) to provide a total photoperiod of 16 h.

2.2 Primer design.

Primers were designed using the Primer3 2.3.4 version integrated in the Geneious software version 8 (Kearse et al. 2012b). Primers were optimised for length (between 18 to 24 bp), G/C content (between 30-80%, with an optimal 50%), an annealing temperature ranging from 54 to 63^{o}C, minimal self or cross compatibility, a maximum melting temperature difference of 5^{o}C between forward and reverse primer, and the presence of a GC clamp at 3' end.

2.3 Nucleic acids extraction.

2.3.1 Genomic DNA extraction.

Unless otherwise specified in each individual chapter, the genomic DNA extraction was performed as follows; first, young leaf tissue was collected in liquid nitrogen and stored at -80^{o}C when not used immediately. Samples were then mechanically ground using either a carbide bead

and a tissue lyser (Qiagen TissueLyser II) or mortar and pestles. Nucleic acids were extracted using 500 µl of 2x Extraction buffer (100 mM Tris-HCl, 1.4M NaCl, 20 mM EDTA, 2% w/v CTAB, 20 mM 2-β-mercaptoethanol, pH 8 with HCl) and incubated for 15 min at 60^{o}C with gentle periodic agitation every 5 min. Solvent extraction was performed twice for optimal purification, using a chloroform-isoamylic alcohol (24:1) solution. DNA was then precipitated with 1mL of Precipitation Buffer (50 mM Tris-HCl, 10mM EDTA, 1% w/v CTAB, pH 8 with HCl) and pelleted by centrifugation for 10 min at 14,000 g. Pellets were resuspended in 300 µl of a 1.5 NaCl solution containing 1 RNase A (25 mg/mL) and incubated at 50^{o}C until total resuspension was achieved. Genomic DNA was precipitated with chilled 95% ethanol and pelleted by centrifugation at 14,000g for 15 min. DNA was finally washed in 70% ethanol, air dried and dissolved in autoclaved Milli-Q water.

2.3.2 RNA extraction and Complementary DNA (cDNA) synthesis.

Tissue samples were harvested, immediately frozen in liquid nitrogen and ground to fine powder using either mortar and pestle or carbide beads and mechanical homogeniser (Qiagen TissueLyserII). RNA was extracted using the Promega SV Total RNA Isolation System (Promega, Madison, WI) according to the manufacturer's instructions. 1µg of total RNA was used to synthesise cDNA with Tetro Reverse Transcriptase (Bioline, London, UK) in a final volume of 20 µL following manufacturer's protocol. Negative control without reverse transcriptase (RT-) was included for all samples to check genomic DNA contamination. cDNA obtained was diluted five times for its final use.

2.4 Polymerase chain reaction (PCR)

2.4.1 Standard PCR

Standard PCR was performed in a final volume of 25 µL containing 50 ng template DNA, 5 µL of 5x reaction buffer, 10 mM dNTPs, 0.2 µM of each primer, 50 mM $MgCl_2$, 0.1 µL of MangoTaq^{TM} DNA polymerase (Bioline, Australia) and autoclaved Milli-Q water to final volume. Reactions were performed in a thermal using the following program: an initial denaturation of 5 min at 94^{o}C, followed by 35-40 cycles (94^{o}C for 40 seconds, annealing temperature for 30 seconds, extension of 1 minute/kb of expected product size) and a final extension of 10 minutes at 72^{o}C.

2.4.2 Colony PCR

Colony PCR was carried out using bacterial colonies suspended in autoclaved Mili –Q water. PCR conditions were the same that those describe in the standard PCR with the following program: initial denaturation at 94°C for 5 minutes followed by 30 cycles (1 minute at 94°C, annealing temperature for 1 minute, extension of 1 minute/kb of PCR product at 72°C) and a final extension of 5 minutes at 72°C.

2.4.3 Real-time polymerase chain reaction (RT-PCR or qPCR)

Gene expression was determined by quantitative PCR (qPCR) using a Rotor-Gene 3000 Real-time Thermal Cycler with Rotor-Gene 6 Version 6.1 (Corbett Research, Australia). Reactions included 2 µL cDNA template, 5 µL 2X SensiFAST SYBR No-ROX mix (Bioline, Australia), 0.4 µM of each primer and autoclaved Milli-Q water to complete a final volume of 10 µL, and were prepared using either a Corbett Robotics CAS-1200TM pipetting robot (Corbett Research, Australia) or a PIRO Pipetting Robot (Lindauer DORNIER GmbH, Germany) with the software provided by supplier.

To check for genomic DNA contamination, housekeeping gene (refer to each chapter for details) was run on the reverse transcriptase negative control (RT-) for each cDNA sample.
Reactions were run for 50 cycles, and all samples were run in duplicate for higher accuracy. As negative and positive controls, a non-template control and standard curve were included in each run. Standard curves were generated from a 10-fold serial dilution from 10^{-1} to 10^{-6} ng/µL. Standard curve regression was considered acceptable if the R^2 value was equal to or higher than 0.99.

Relative expression to reference gene was calculated on the basis of non-equal amplification efficiencies and deviation in threshold cycle using the means for two technical replicates (Pfaffl 2001).

2.4.4 Visualization of nucleic acids

To visualise PCR products and check DNA/RNA integrity, samples were separated by electrophoresis on agarose gel in TAE buffer (40mM Tris Acetate and 1mM EDTA), stained

with GoldView™ Nucleic Acid Stain (Acridine orange; SBS Genetech, China) and visualised under UV light. Estimation of PCR products size was done using an appropriate DNA ladder.

2.4.5 Purification of PCR products

PCR products were purified using Promega Wizard® SV Gel and PCR Clean-Up System (Promega, USA) and eluted in sterile, nuclease free water following manufacturer's instructions.

2.5 Cloning of PCR products

Purified PCR products were ligated into pGEM®-T Easy vectors (Promega, USA), in accordance with the manufacturer's instructions. Competent *E. coli* cells were transformed by electroporation at 1200 V and allowed to recover in 400 µL of Luria Broth (LB, 10 g/L Bacto-tryptone, 5 g/L Bacto-yeast extract, 10 g/L NaCl, pH 7.5) with incubation at 37°C for 1 hour with shaking. Transformation reactions were spread across LB agar plates (ingredients as for LB broth with 15 g/L agar and 100µg/mL ampicillin and 1 µL/mL X-gal added) and incubated overnight at 37°C. White colonies were screened for the desired length insert by colony PCR.

2.6 Nucleic acids quantification

Concentration of DNA, RNA and PCR products was determined using a NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific, USA) according to manufacturer's instructions.

2.7 Sequencing and sequence analysis

Purified PCR products were sent for sequencing to Macrogen Inc. (Seoul, Korea). Sequences were inspected, edited and annotated using Geneious software version 8 (Kearse et al. 2012b). Sequence identity was confirmed by BLAST search or alignment with existing sequence.

2.8 Molecular markers design and genotyping

All markers in this thesis are HRM markers. Primers were designed to amplify small fragments (<200 bp) targeting InDels and SNPs (except A/T or G/C SNPs). HRM markers were tested and scored in segregating populations using a Rotorgene Q HRM machine (Qiagen). Reactions were

prepared using either a Corbett Robotics CAS-1200TM pipetting robot (Corbett Research, Australia) or a PIRO Pipetting Robot (Lindauer DORNIER GmbH, Germany) with the software provided by supplier, and included 50 ng template, 0.5 µM of each primer, 7.5 µL SensiFAST™ HRM Mix (from SensiFAST™ HRM Kit, Bioline), and sterile milli-Q water to complete 15 µL. HRM reactions were performed following this program: 95°C for 5 minutes, 50 cycles [95°C for 10 seconds, annealing temperature (Tm; 50-60°C) for 30 seconds], 95°C for 5 minutes, 50°C for 5 minutes, HRM (temperature increasing with 0.1°C increments from 60-90°C, or from product melt temperature -5°C to +5°C). HRM results were analysed with Rotor-Gene® ScreenClust HRM® Software (Qiagen).

2.9 Online resources.

All sequences used in this study for gene identification, homology analysis and primer design were obtained from the sources shown in table 2.1.

Table 2.1 Detail of online resources used for sequence information

Species	Website	Version	Reference
Cicer arietinum			
Genome	GenBank (www.ncbi.nlm.nih.gov)	1.0	(Varshney et al. 2013c) (Jain et al. 2013)
Markers	https://www.integratedbreeding.net/98/communities/genomics-crop-info/agricultural-genomics/markers/ https://www.coolseasonfoodlegume.org/		
Arabidopsis thaliana	The Arabidopsis Information Resource (www.arabidopsis.org)	TAIR10	(Berardini et al. 2015)
Medicago truncatula	Medicago truncatula Genome Database (www.medicagogenome.org)	4.0	(Krishnakumar et al. 2015)
Pisum sativum	The Pea RNA-Seq gene atlas (bios.dijon.inra.fr/FATAL/cgi/pscam.cgi)	-	-
Glycine max	Phytozome (phytozome.jgi.doe.gov)	2.0	(Goodstein et al. 2012)
Phaseolus vulgaris		2.1	
Lens culinaris	KnowPulse (knowpulse2.usask.ca)	1.2	-

2.10 Software

Information about the versions of the informatics software described here and further details can be found in table 2.2.

Sequence editing and alignments were performed in Geneious 8 using three different aligners available as plugins; MUSCLE (Edgar 2004), ClustalW (Thompson et al. 1994) and MAFFT (Katoh et al. 2002). Phylogenetic trees were performed using either PAUP* (when maximum parsimony used to build the tree) MEGA 6 (maximum likelihood). Alignments for figures were graphed with GeneDoc software.

All statistical analyses were conducted using either GraphPad Prism or IBM SPSS statistics version 22, with a significance level of 0.05 for all analyses unless otherwise specified.

Genetic maps were performed using Join Map 4 and graphed using either this software or MapChart. QTL analysis were done with MapQTL 6.0.
Gene expression and phenotype charts were made using GraphPad Prism. If necessary, pictures, graphs and phylogenetic trees were edited to make the final figures in Adobe Illustrator.

Table 2.2 Information about software packages used

Software	Version	Developer and Website	Reference
GraphPad Prism	6.01	GraphPad Software (www.graphpad.com)	-
Geneious	8.1.9	Biomatters (www.geneious.com)	(Kearse et al. 2012a)
MEGA	7.0.20	MEGA Software (www.megasoftware.net)	(Kumar et al. 2016)
MapChart	2.30	Wageningen Plant Research (www.wur.nl/en/show/Mapchart-2.30)	(Voorrips 2002)
JoinMap	4	Kyazma B.V. (www.kyazma.nl)	(Stam 1993)
MapQTL	6	Kyazma B.V. (www.kyazma.nl)	(Van Ooijen and Maliepaard 1996)
Adobe Illustrator	cs5	Adobe Systems www.adobe.com/Illustrator	-
GeneDoc	2.7.0	www.psc.edu/biomed/genedoc	(Nicholas and Nicholas 1997)
IBM SPSS Statistics	22.0	IBM Corporation (www.ibm.com/analytics/us/en/ technology/spss)	-
PAUP*	4.0b10	Smithsonian Institution (paup.sc.fsu.edu)	(Swofford 2001)

Chapter 3. Conservation of Flowering Genes in chickpea

3.1 Introduction

Legume crops are tremendously important on a global scale, providing food for humans and livestock, and enhancing soil fertility through symbiotic nitrogen fixation, but until recently were considered to be genomic "orphans" due to a lack of genetic and genomic tools. As result, crop improvement in these species has largely occurred through traditional breeding approaches, with limited or no impact from molecular technologies (Varshney et al. 2009a). In the last two decades, however, important genomic advances have been made in a great number of legume species, opening up the possibility of accelerating legume crop improvement through application of genomics-assisted breeding.

In the particular case of chickpea, the transition to the genomic era is now well-established. As outlined in Chapter 1 (see section 1.6), the development of genetic and physical maps began in the nineties and progressed with the incorporation of new types of markers that allowed researchers to compare different maps, unify linkage groups nomenclature, and facilitated the study of chickpea synteny with other legumes. (Zhang et al. 2010; Pfaff and Kahl 2003; Hiremath et al. 2012; Nayak et al. 2010; Palomino et al. 2009; Winter et al. 1999; Winter et al. 2000; Thudi et al. 2011; Cobos et al. 2005; Millan et al. 2010; Deokar et al. 2014; Tekeoglu et al. 2002). It is now apparent that despite significant rearrangements, genomic structure is well conserved across legume species (Nayak et al. 2010; Lee et al. 2017; Gujaria-Verma et al. 2014), which has great significance since it allows the transfer of knowledge obtained in other legumes (e.g. genome structure, position of potential genes, potential identity of genetic loci) for comparative studies to chickpea. In this respect, the legume model species *Medicago truncatula* is of particular relevance because among all the sequenced legume genomes it is taxonomically the closest to chickpea and its synteny with other legume species is well-documented, enabling cross-species comparison (Tang et al. 2014; Choi et al. 2004). Coupled to the development of molecular markers comes the possibility of linking them to valuable traits, as reflected by the considerable number of Quantitative Trait Loci (QTL) analyses published for chickpea in the last two decades, linking different regions of chickpea genome with the control of agronomically interesting traits (Millan et al. 2014).

Without any doubt, the definitive genomic tool for any species is the sequencing of its whole genome, since it enables the identification of genes and functional elements responsible for expression of phenotype and provides a platform for gene mapping, gene isolation and molecular breeding (Varshney et al. 2013a). The sequencing of the chickpea genome from representative *desi* and *kabuli* types was reported in 2013 (Jain et al. 2013; Varshney et al. 2013c), providing a reference that has facilitated the re-sequencing of a wider range of varieties (Thudi et al. 2016; Li et al. 2017; Das et al. 2015a).

Another consequence of the genome sequence availability is that it enables sequence-based markers from older studies to be located on the physical map. This allows the chromosomal location of previously reported QTLs to be physically defined, and makes it possible to identify potential (candidates) genes governing a trait. Since molecular research in chickpea is restricted to recent years, it is necessary to look for available information in other species with advanced molecular knowledge in order to find these candidates genes. In the particular case of flowering control, much of our current understanding is derived from studies using the model species *A. thaliana*. Hundreds of genes have been described affecting this trait as response to external and endogenous signals, and the general model described in this species is well conserved among plants, as reflected by the fact that many orthologs of these genes maintain their capacity to alter time to flower in even far related species. Although still far from the level reached in Arabidopsis, the molecular physiology of flowering time is increasingly well studied in the legume species pea, soybean and Medicago, as reviewed in (Jung et al. 2012; Hecht et al. 2005; Weller and Ortega 2015; Kim et al. 2013b; Kim et al. 2012).

In chickpea more than 50 QTLs have been reported for flowering time, distributed across all eight chickpea linkage groups (LG) (table 1.2). Multiple reports of QTL on chromosomes 3 and 4 indicate the existence of two genomic regions of particular importance for chickpea flowering time variation. Of special relevance is the central portion of LG3 where reports from several different inter and intraspecific populations suggest the presence of a major gene (Cobos et al. 2009; Mallikarjuna et al. 2017; Hossain et al. 2010; Rehman et al. 2011; Aryamanesh et al. 2010). Moreover, the syntenic genomic region has been associated with flowering control in several other legumes including *Medicago truncatula* (Pierre et al. 2008), *Vicia faba* (Cruz-Izquierdo et al. 2012), *Lupinus angustifolius* (Nelson et al. 2006), *Lotus japonicus* (Gondo et al. 2007), and

Medicago sativa (Robins et al. 2007), suggesting the possible involvement of a major flowering gene with conserved function across these different species. This region is known to harbour a cluster of *FT* homologs (*FTa1*, *FTa2* and *FTc* in Medicago), of which the *FTa1* gene has been shown to have a key role in regulation of flowering time and vernalization response in pea and Medicago (Weller and Ortega 2015; Laurie et al. 2011; Hecht et al. 2011). These observations have led to the suggestion that genes in this cluster represent the most likely candidates for the conserved QTL across different legume species, and in particular for the major recurrent QTL in chickpea (Weller and Ortega 2015).

However, despite the likely importance of this *FT* cluster, the existence of other potential candidate genes in the region remains to be thoroughly explored. In order to examine this in more detail, and to identify candidates for other previously-described regions, this chapter aims to (a) investigate the conservation of *A. thaliana* flowering-related genes in chickpea (b) anchor previously-described flowering time QTL to the physical map, through analysis of marker sequences (c) examine the co-location of flowering-related genes in relation to these QTLs in order to identify the most plausible candidate genes in each case and (d) confirm the colocation and likelihood of the *FT* cluster as the most promising candidate for the recurrent QTL interval in LG3 abovementioned, which has been speculated about but not clearly demonstrated.

3.2 Materials and methods

Information about the software used for the analysis (including version and references) can be found in General Materials and Methods, section 2.10.

Sequence databases

The genomes from the chickpea *kabuli* cultivar CDC Frontier (Varshney et al. (2013c), BioProject PRJNA190909) and the *desi* accession ICC4958 ((Jain et al. 2013), BioProject PRJNA7895) were obtained from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov) and uploaded into Geneious 8 software. Protein sequences and accession numbers from Arabidopsis and other legumes, *Medicago truncatula*, *Pisum sativum* (pea), *Phaseolus vulgaris* (common bean) and Glycine max (soybean) were obtained from the databases specified in Section 2.9 and also uploaded into Geneious.

Mapping of markers in chickpea genome

To determine the physical position of genetic markers previously shown to be linked to flowering time QTLs, relevant primer sequences were acquired from the corresponding studies and mapped in the two *Cicer* genomes (CDC Frontier and ICC4958) using the default Geneious mapper and custom parameters: maximum gap size of 4 or 20% of primer length; maximum ambiguity, 4; maximum mismatches per read, 20%; word length, 7.

Identification and phylogenetic analysis of chickpea flowering-related genes

Protein sequences from Arabidopsis flowering genes and, when available, homologs from Medicago and pea were used as query to search for chickpea homologs by BLASTp (Basic Local Alignment Search Tool) in NCBI database with default parameters. To retrieve only chickpea proteins, *Cicer arietinum* (taxid:3827) was selected in the 'Organism' BLAST option.

To establish the most plausible identity of the chickpea sequences obtained by BLASTp, protein sequences were aligned with those from Arabidopsis and other legumes in Geneious 8 using 3 different alignment algorithms; ClustalW, MAFFT and MUSCLE using default parameters in each case. The alignment with a higher score in the BLOSUM62 substitution matrix was used to

construct a maximum parsimony or maximum likelihood tree in order to assess the phylogenetic relationships among the genes retrieved, and named them accordingly.

Figure construction

The physical positions of all chickpea genes of interest were obtained from NCBI using the genome of cultivar CDC Frontier as reference. Finally, this information together with that obtained from the markers was used to draw a chickpea genome map of flowering candidate-genes and reported flowering-time QTLs. The figure was graphed using MapChart.

3.3 Results

3.3.1 Candidate genes in chickpea genome

In order to find chickpea flowering candidate genes, 234 *Arabidopsis thaliana* genes were selected based on their known role on flowering control or their close phylogenetic relationship to such genes. BLAST searches in chickpea yielded a total of 255 potential genes, indicating overall that the majority of the genes used in the query have at least one close chickpea homolog. Since the conservation of flowering genes in other legumes has already been discussed in previous studies (Jung et al. 2012; Hecht et al. 2005; Weller and Ortega 2015; Kim et al. 2013b; Kim et al. 2012), we will mention here only gene families that are key to the flowering network, such as the *CONSTANS-LIKE* (*COL*) and the PEBP gene family of *FT*-like genes. A list of all Arabidopsis genes used as query and their accession numbers is available in appendix 3.1, together with references to further information about chickpea related sequences (appendices 3.2 to 3.25).

MADS-box family

MADS-box transcription factor proteins have a well-known role in many developmental processes in plants, particularly during reproductive development. The MADS-box family is a huge group of proteins, with more than 100 members in Arabidopsis. They can be divided in two broad groups called Type I and Type II (also known as MIKC) based on major structural differences. For example, MIKC-type are so called because the MADS domain common to all family is followed by an Intervening (I), a Keratin-like (K) and a C-terminal domain (Gramzow and Theißen 2013). Only members of this group are reported to be involved in flowering regulation and are therefore relevant to this study. Homologs for almost all MIKC-MADS proteins were found in chickpea, although the most notable exception was the floral repressor *FLOWERING LOCUS C* (*FLC*) and the small group of proteins in the same sub-family known as *MADS AFFECTING FLOWERING* (*MAFs*), which have a key role in vernalization response in Arabidopsis (Ratcliffe et al. 2003). Despite recent work suggesting that the *FLC* clade could have originated early in the diversification of the eudicots, their function outside the *Brassicaceae* is still uncertain (Perilleux et al. 2013; Reeves et al. 2007; Vogt et al. 2014). The genetics underlying vernalization in other systems (e.g. cereals) points that this response can be

quite divergent across plant kingdom (Alexandre and Hennig 2008; Vogt et al. 2014). Putting together this degree of variation and the fact that previous comparative studies in legumes failed to find any convincing evidence for the *FLC* clade (Kim et al. 2013b; Hecht et al. 2005), the absence of these genes in chickpea can be then considered as an expected result.

The analysis also confirmed the duplication of some important MADS-box genes in chickpea and other legumes relative to Arabidopsis, including *FRUITFULL* (*FUL*), *PISTILLATA* (*PI*), *AGAMOUS* (*AG*) and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), which are represented with at least 2 homologs in chickpea and other legumes (for a tree of the MADS-MIKC, refer to appendix 3.2)

Photoreceptors and light signalling

Light is for most plants the main environmental regulator of flowering. Information about photoperiod is perceived in the leaves through a series of photoreceptors well documented in Arabidopsis, including *PHYTOCHROMES* (*PHY*) and *CRYPTOCHROMES* (*CRY*) which are among the first to be discovered (Mathews 2006; Lin and Todo 2005). More recent studies have uncovered the identity of other light-interacting proteins with an impact on flowering time, such as *ZEITLUPE* (*ZTL*) or *FLAVIN-BINDING, KELCH REPEAT, F-BOX 1* (*FKF1*) (Kim et al. 2007; Imaizumi et al. 2003). The conservation of genes belonging to this category has been studied in other legumes, founding an overall high degree of conservation compared to Arabidopsis (see Chapter 1, Section 1.5.2). As in Medicago, most of the photoreceptors included in the BLAST search have a chickpea homolog (appendix 3.3), except for the remarkable absence of *PHYC* and the existence of two *PHYA* homologs. While the lack of *PHYC* is common to all legumes, the duplication of chickpea *PHYA* is a novel result, as it has been documented in legumes but it was believed to be restricted to the phasioloid clade of legumes (Weller and Ortega 2015).

Photoperiod and Circadian Clock genes

Circadian rhythms are the temporal oscillation of genetic, metabolic and physiological processes, allowing organisms to anticipate day–night changes in the environment. In plants, photoperiod and circadian clock are deeply interconnected in a complex molecular mechanism that integrates environmental and endogenous cues to regulate a wide range of responses, including flowering

time (Nakamichi 2011; Greenham and McClung 2015). The molecular composition of Arabidopsis circadian clock is well studied, consisting of a series of interlocked feedback loops involving transcriptional, post-transcriptional and post-translational regulation of more than 25 genes (McClung 2014; Huang and Nusinow 2016). In this species, clock influences flowering time through the regulation of the key flowering gene *CONSTANS* (*CO*). Together with light signalling, circadian clock regulates the levels of CO protein to ensure it only accumulates at the appropriate time of the year (Suárez-López et al. 2001; Shim and Imaizumi 2015).

In legumes, orthologs of the genes involved in photoperiod and circadian clock are well conserved according to previous reports (Ridge et al. 2016; Hecht et al. 2005; Jung et al. 2012; Kim et al. 2013b). The results obtained here show a high degree of conservation in the number and sequence homology between chickpea proteins and those from other legumes (appendices 3.4 to 3.10). No homolog of *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*) was found in chickpea. This is an essential gene involved in maintenance of circadian rhythms in Arabidopsis whose absence it is worth mention (Mizoguchi et al. 2002). In any case, previous reports in Medicago and soybean coincide in the same result (Hecht et al. 2005), suggesting that the lack of *CCA1* could be a conserved feature among legumes.

Phase change and Autonomous pathways

In Arabidopsis, the "autonomous" flowering pathway has been used as general term to group all those endogenous factors with effects on flowering that are independent from environmental cues. These include factors such as hormones, microRNAs involved in age and phase transitions, or genes influencing the expression of other floral integrators. No major differences were found in the conservation of the autonomous pathway genes in chickpea (appendix 3.11); most of the Arabidopsis genes had a chickpea (and Medicago) homolog. Since many of these genes are involved in the regulation of *FLC* in Arabidopsis (e.g. FCA, FLD, FPA and FLK), but this target is not present in chickpea (or any legume, as above stated), the mechanism of action of such genes in legumes could have diversified and is still to be determined (Cheng et al. 2017). The conservation of other genes analysed in this category, such as those involved in phase change pathway (appendixes 3.12 to 3.16**)** or gibberellin oxidases (appendix 3.17) was overall very high; chickpea homologs for most of these genes were found, and they show a high level of homology with those from other legume species.

Vernalization pathway

In Arabidopsis, vernalization response is mostly mediated through the MADS-box *FLC* and its transcriptional activator *FRIGIDA (FRI)* (Michaels and Amasino 1999, 2001). *FLC* is also the target of many other genes involved in vernalization processes in this species. Such genes, despite the lack of their target, were found to be conserved in legumes, including chickpea in the present study (appendix 3.18). An interesting observation is that 3 genes were found related to *VERNALIZATION INSENSITIVE3* (*VIN3*), contrasting with the lack of homologs reported in soybean (Jung et al. 2012).

CONSTANS and CONSTANS-like family

CONSTANS (CO) was the first member discovered in a plant family of zinc finger proteins that rapidly expanded to 17 genes named *CONSTANS-LIKE* (*COL*) genes (Putterill et al. 1995b; Lagercrantz and Axelsson 2000). Arabidopsis CO protein contains two B-box zinc fingers close to the amino terminus, a plant specific CCT (CO, CO-Like, TOC1) domain near the carboxyl terminus and a six-residue T motif at the carboxyl end (Griffiths et al. 2003; Strayer et al. 2000; Putterill et al. 1995a). The variation in these features subdivide the *COL* family in three groups; type I genes (*AtCO* and *AtCOL1* to *AtCOL5)* include two B-boxes and one CCT domain. Type II *COL* genes *(AtCOL*6 to *AtCOL8* and *AtCOL16)* present only one B-box and one CCT domain, while type III (*AtCOL9* to *AtCOL15)* have one full B-box, a second diverged zinc finger and one CCT domain (Robson et al. 2001).

In Arabidopsis, *CO* has a central role in the photoperiod response mechanism. Signals from photoreceptors and circadian clock are integrated to enhance *CO* mRNA and stabilize CO protein only under long day conditions. Once accumulated, CO promotes the expression of the floral integrators *FT*, *TSF* and *SOC1* to ensure a correct timing of floral induction (Suárez-López et al. 2001; Kim et al. 2008; Searle and Coupland 2004). Even though an important degree of variation have been reported in their mechanism of action, this flowering-promoter character seems to be well conserved in *CO* orthologs among distantly related species, including potato (Martinez-Garcia et al. 2002), rice (Yano et al. 2000), green algae (Serrano et al. 2009) and trees (Bohlenius et al. 2006). Less is known about the function of other genes in the broader *COL* family. Despite their high homology with *CO*, *AtCOL1* and *AtCOL2* causes no changes in

flowering time (Ledger et al. 2001). The overexpression of *AtCOL5* can induce flowering under short day conditions but, conversely, *col5* mutants show no alteration in flowering time, suggesting either a lack of biological function in wild-type plants or a redundant role with other flowering genes (Hassidim et al. 2009). *AtCOL3* is involved in root development and, like *AtCOL9* and rice *OsCOL10*, can repress flowering (Datta et al. 2006; Cheng and Wang 2005; Tan et al. 2016), indicating a functional divergence of this family.

Eleven *CONSTANS-like* homologs (*CaCOLs*) were found in chickpea genome (appendix 3.19), a number similar to those described in pea and Medicago (Ridge et al. 2016; Wong et al. 2014). To clarify their identities, related *COL* sequences from Arabidopsis, soybean (Wu et al. 2014), Medicago and pea were used to build a phylogenetic tree (Fig 3.1). Each *CaCOL* was then named based on its homology with *COLs* from other legumes and following the nomenclature adopted byWong et al. (2014). Despite a reduction in the number of genes in legumes (11) compared to Arabidopsis (17), homologs in each of the 3 groups (types I, II and III) were found in all 4 legumes analyzed. The tree obtained is also consistent with that described in legumes (Wu et al. 2014; Wong et al. 2014) (Fig 3.1).

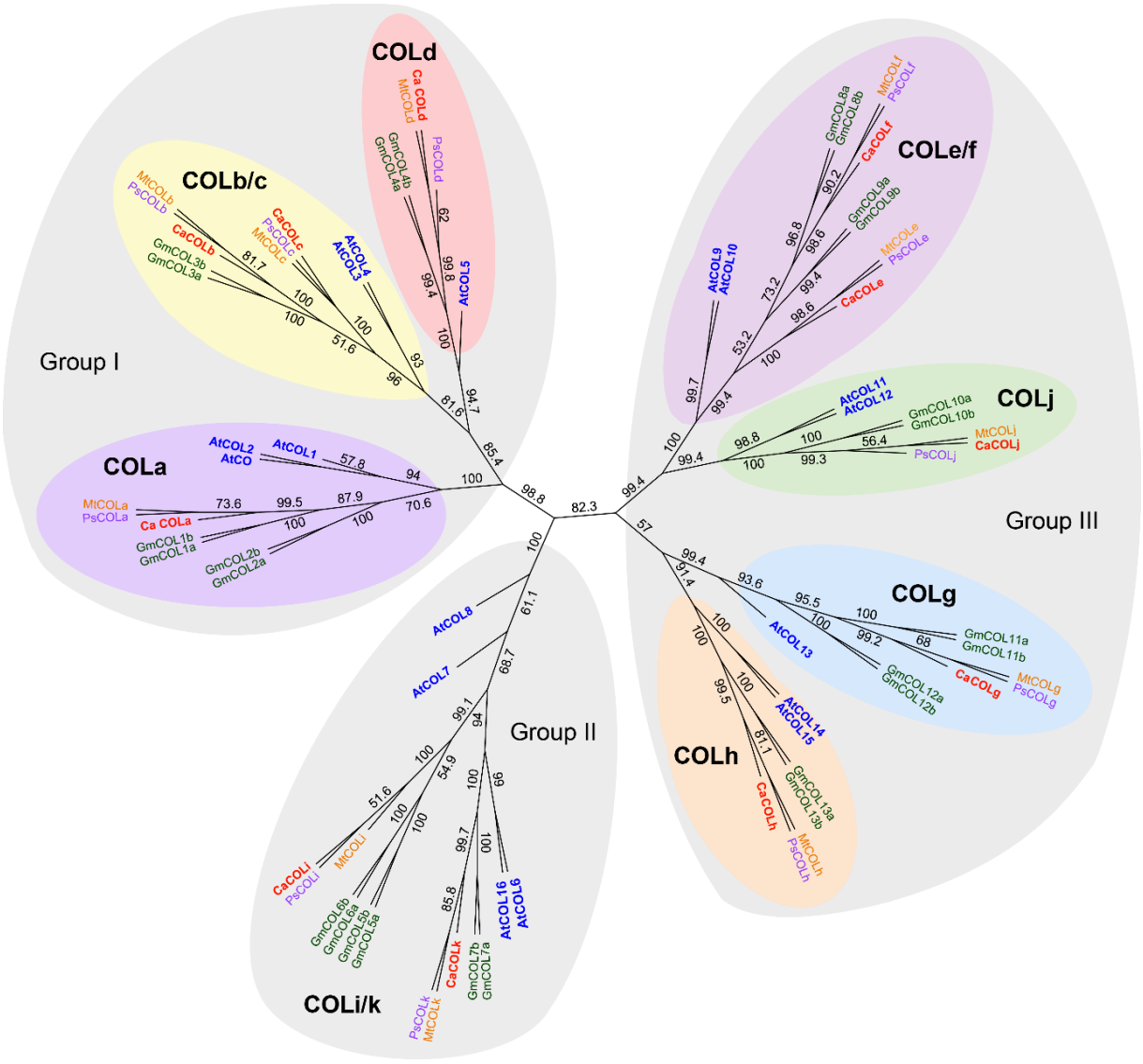


Figure 3.1 Maximum parsimony tree derived from the alignment of *CONSTANS-LIKE* genes (*COL*) in *Arabidopsis thaliana* (At, blue), *Cicer arietinum* (Ca, red), *Glycine max* (Gm, green), *Pisum sativum* (Ps, purple) and *Medicago truncatula* (Mt, orange). Protein sequences of accessions listed in appendix 3.19 were aligned with MUSCLE and the tree was build using PAUP* in Geneious 8 software. Numbers in branches represent bootstrap support from 1000 replications.

Characterization of the chickpea PEBP family

The phosphatidylethanolamine-binding protein (PEBP) family is a highly conserved group of proteins present in all eukaryote kingdoms. In Arabidopsis, *FLOWERING LOCUS T* (*FT*) and *TERMINAL FLOWER 1* (*TFL1*) are the most famous members of this family that includes four more genes, namely *TWIN SISTER OF FT* (*TSF*), *MOTHER OF FT AND TFL1* (*MFT*), *BROTHER OF FT AND TFL1* (*BFT*) and *ARABIDOPSIS THALIANA CENTRORADIALIS* (*ACT*

or *CEN*). PEBP homologs have been extensively studied in a wide range of plants, grouping them in four subclades: *FT* (and *TSF*), *MFT*, *TFL1* (including *ACT*) and *BFT* (Danilevskaya et al. 2008; Kikuchi et al. 2009). In plants, they have well-known roles in the induction of flowering and the regulation of plant architecture (Karlgren et al. 2011). The *FT* gene is the main integrator of the different flowering pathways governing flowering in Arabidopsis. Once expressed, FT protein acts as a florigen that moves from the leaves to the shoot apex, where it binds to the bZIP transcription factor FD to form a flowering inducing complex (Corbesier et al. 2007; Turck et al. 2008). This general role appears widely conserved across all plants, as proved by the fact that *FT* paralogues have been reported to promote floral transition in a wide range of plants, including cereals (Komiya et al. 2008; Meng et al. 2011), woody perennials species such as poplar (Hsu et al. 2006), apple (Kotoda et al. 2010), kiwifruit (Voogd et al. 2017), cassava (Bull et al. 2017) and vine grapevine (Carmona et al. 2007). By contrast, *TFL1* is a floral repressor that delays transition from vegetative to reproductive state. This antagonistic role to *FT* is also well conserved in many species ranging from herbaceous species to woody perennial, and it is exerted through repression of several genes downstream of *FT* such as *LEAFY* (*LFY*) and *APETALA1* (*AP1*), but also perhaps by competing with FT protein to bind FD, with which it weakly interacts (Wickland and Hanzawa 2015).

The knowledge of this family in legume species has improved dramatically in recent years. Several studies in pea, Medicago, *Lotus japonicus* and soybean have characterized the PEBP family at the molecular level, with a focus on the *FT* genes. Several *FT* and *TFL1* homologs can be found in most species and phylogenetic analysis suggests relatively old diversification within legumes, resulting in an expansion of the family. In temperate legumes, 3 subgroups within *TFL1* have been reported, namely *TFL1a*, *TFLb* and *TFL1c*. In a similar manner, 3 subgroups (*FTa*, *FTb* and *FTc*) can be distinguished within the *FT* clade, and unlike the previous case this division seems to be valid to all galegoid and phaseoloid legumes (Hecht et al. 2011; Takeshima et al. 2016; Yamashino et al. 2013; Kong et al. 2010; Ono et al. 2010; Sun et al. 2011; Wang et al. 2015).

Twelve genes were identified as PEBPs by BLASTp search against the chickpea genome. Phylogenetic analysis with protein sequences from Arabidopsis and the legume species Medicago, pea and soybean grouped them in four different sub-families, consistent with those

abovementioned (Fig 3.2). Only one homolog was found in chickpea for each *MFT* and *BFT* genes, according to expectations. However, the composition of the *FT* and *TFL1* subfamilies in chickpea showed some unique features relative to other legumes so far reported. In chickpea, one of each *TFL1a* and *TFL1b* were found, whereas three different *TFL1c* proteins were retrieved and named *CaTFL1c1* (GeneID 101495644), *CaTFL1c2* (GeneID 101491943) and *CaTFL1c3* (GeneID 101492277). An alignment of the three proteins can be found in appendix 3.20. *CaTFL1c1* and *CaTFL1c2* show a high homology (98.9%), differing only in two residues, while *CaTFL1c3* is more divergent lacks the N-terminus third of the protein, and is therefore probably a pseudogene. *CaTFLc2* and *CaTFL1c3* genes are located in tandem on chromosome 8 of the reference genome assembly (cultivar CDC Frontier), whereas *CaTFL1c1* was found in an unplaced scaffold. Since none of these genes could be mapped in the other available chickpea genome (accession ICC4958), this duplication must be considered with caution until further confirmation proves it real.

As previously mentioned, three *FT* subclades can be found in legumes, and a common nomenclature has been proposed for them that was followed here (Hecht et al. 2011). Consistent with this, five chickpea genes were classified as *FT-like* genes: three corresponded to the *FTa* subclade (named *CaFTa1*, *CaFTa2* and Ca*FTa3*), and one to each *FTb* (*CaFTb*) and *FTc* (*CaFTc*). As reported in Medicago and pea, *CaFTa1*, *CaFTa2* and *CaFTc* are located together in a consecutive position on chromosome 3 (Laurie et al. 2011; Hecht et al. 2011). A novel *FT* homolog has been identified in the present study belonging to the *FTa* clade, so it was named *CaFTa3*. It is not a duplication restricted to chickpea as homologs of this gene were also found in Medicago, pea, soybean and other legumes of different clades (Książkiewicz et al. 2016; Nelson et al. 2017). The only difference in the chickpea *FT* family compared to other legumes is the apparent presence of only one *FTb* gene. Medicago and pea each possess 2 highly similar *FTb* paralogs, which are located in tandem at the bottom of the chromosome 7 in Medicago (Laurie et al., 2011) and in the syntenic region of linkage group 5 in pea (Hecht et al., 2011). Only one *FTb* homolog was found on chickpea chromosome 2, a position that does not fit the synteny with the other two species, as this gene would be expected to be in a proximal region of chromosome 3. This could indicate a translocation event at some point of the recent chickpea evolution that could also be coupled with the loss of one of the genes or, more likely, that pea and Medicago

duplications are more recent events and that chickpea *FTb* could represent the single ancestral gene.

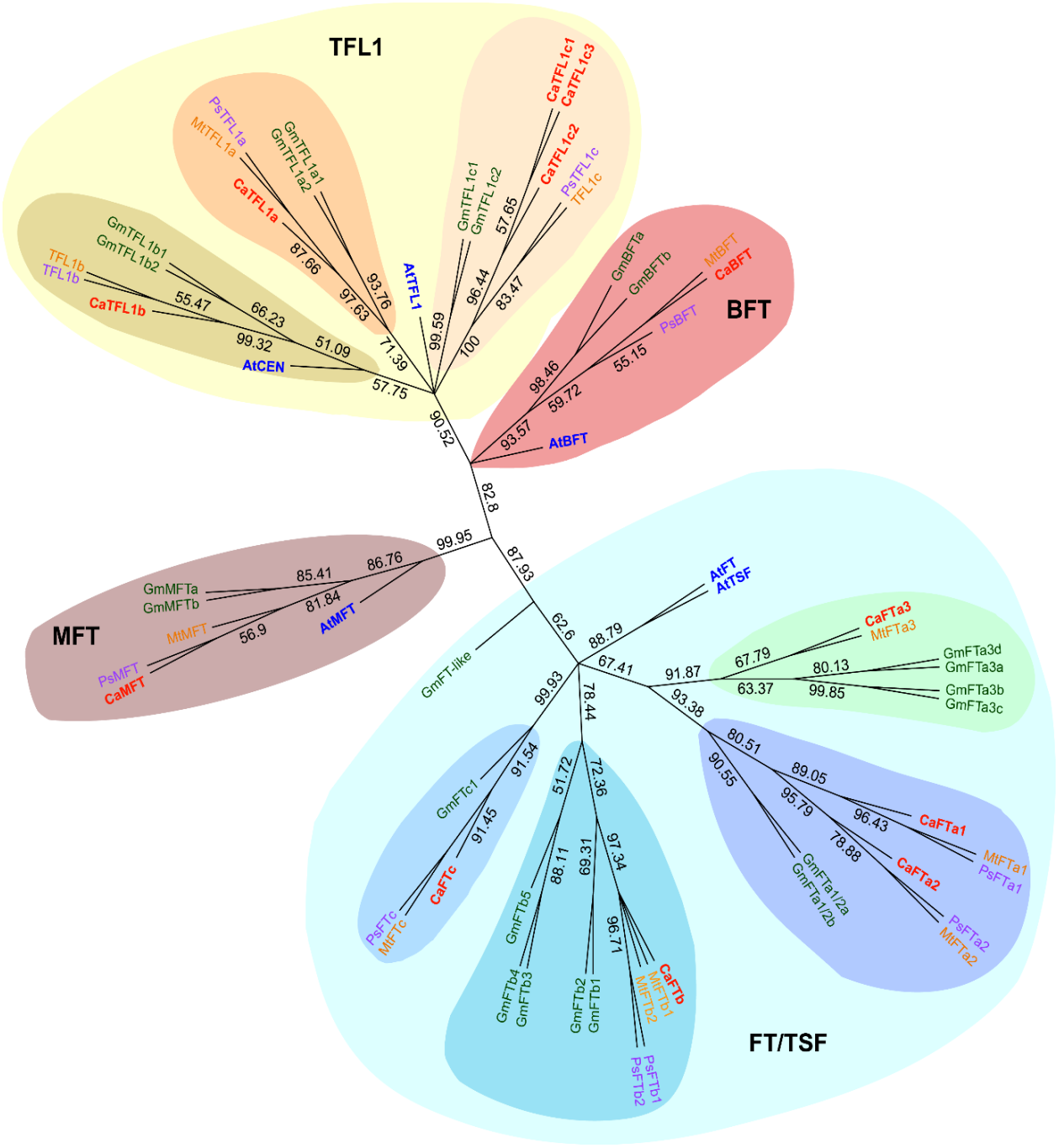


Figure 3.2 Maximum parsimony tree derived from the alignment of PEBP proteins in *Arabidopsis thaliana* (At, blue), *Cicer arietinum* (Ca, red), *Glycine max* (Gm, green), *Pisum sativum* (Ps, purple) and *Medicago truncatula* (Mt, orange). Accession number of the genes, alignment and homology ratios can be found in appendices 3.21 to 3.25. Sequences were aligned with MAFFT and the tree was build using PAUP* in Geneious 8 software. Numbers in branches represent bootstrap support from 1000 replications.

The genomic structure of the chickpea *FT* genes is presented in Figure 3.3, consisting in 4 largely conserved exons and three introns more variable in size, consistent with those described in other species (Wickland and Hanzawa 2015; Wang et al. 2015; Danilevskaya et al. 2008; Hecht et al. 2011; Laurie et al. 2011; Carmona et al. 2007). The size of the genes ranged from 2898 bp (*FTb*) to 16510 bp (*FTa2*), a variation mainly due to intron sizes, which varied from 93 bp in *FTb* to 8.1 kb in *FTa2*. Interestingly, while most *PEBP* genes in plants commonly have only 4 exons, *CaFTa2* is annotated in NCBI (GeneID 101496618) as having a much longer 5' UTR containing three additional exons and introns (Figure 3.3). RNAseq data available in the same NCBI entry provides evidence that these extra exons are transcribed, but the functional implications of this unique structure still need to be determined.

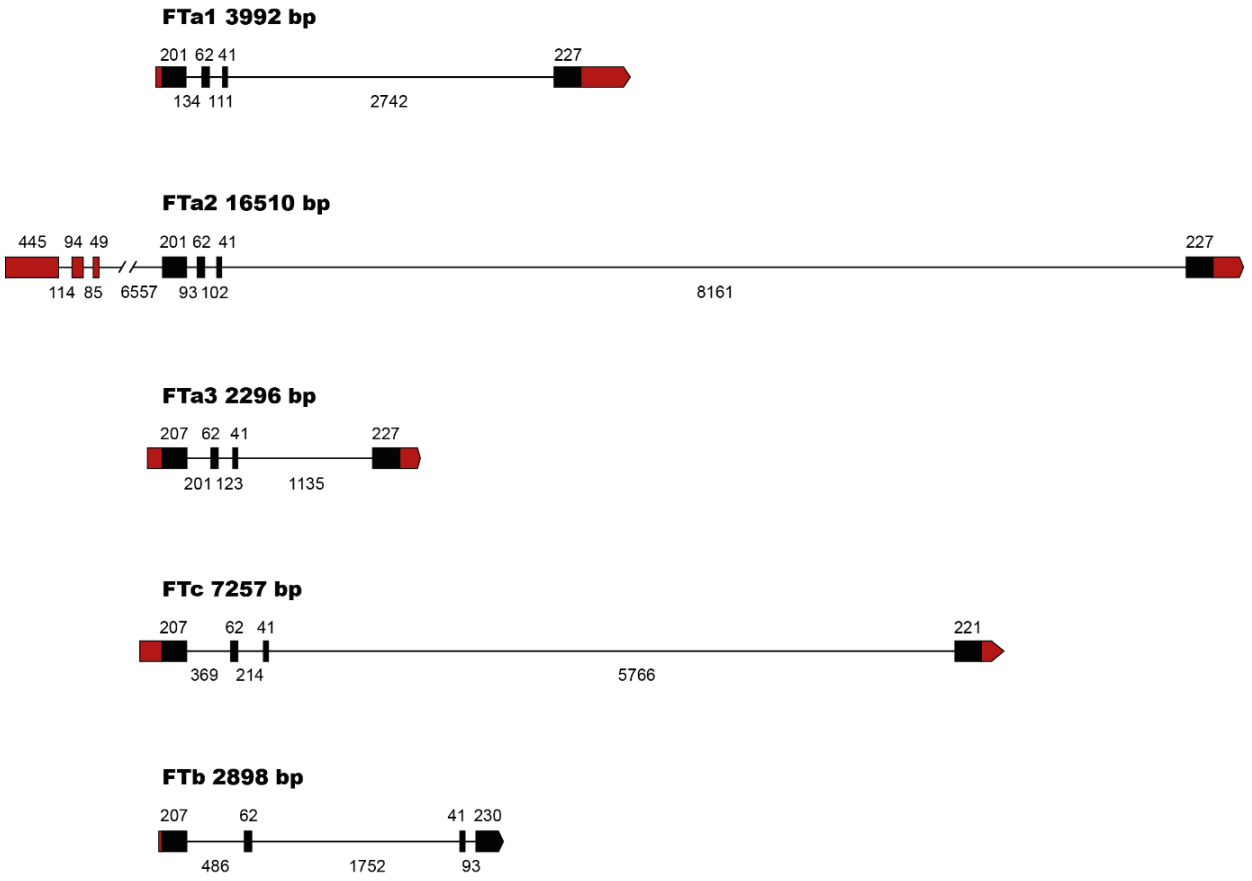


Figure 3.3 Graphic representation of the genomic sequences of chickpea *FT* genes. Black boxes indicate exons (sizes over the line, in bp), red boxes mRNA and lines correspond to introns (sizes under the line).

As shown in Figure 3.4, amino acid conservation was high between the FT proteins from chickpea and their homologs in other species, especially with the taxonomically closer legumes

Medicago and pea. Previous studies have identified several amino acids contributing to an opposite effect in FT and TFL1. A single amino acid substitution (Tyr85 in FT and His88 in TFL1) is enough to revert the function of these proteins (Hanzawa et al. 2005). Specific mutations in four other residues (Glu109, Trp138, Gln140, and Asn152) can have similar effect (Ho and Weigel 2014). Finally, 14 amino-acids within the fourth exon (the so called Segment B) are highly conserved among FT orthologs, as they form an external loop essential for the correct interactions of this protein (Ahn et al. 2006). The conservation of these key residues was evaluated; with the exception of Gln140 and Asn152, these amino acids are conserved in all chickpea *FTs* (highlighted in red in Fig 3.4). The substitution Asn152Asp was found in *CaFTa2*. Although Asn152 is a conserved residue located in the external loop, this change was largely tolerated according to mutagenesis studies realized by Ho and Weigel (2014). In CaFTc protein, Gln140 was replaced with His. This change is not particular of chickpea but present in all FTc proteins in legumes, as previously reported by Hecht et al. (2011).

Overall, there is little evidence from protein sequence to suggest that chickpea FTs differ in any major way from their orthologs in other legume species.

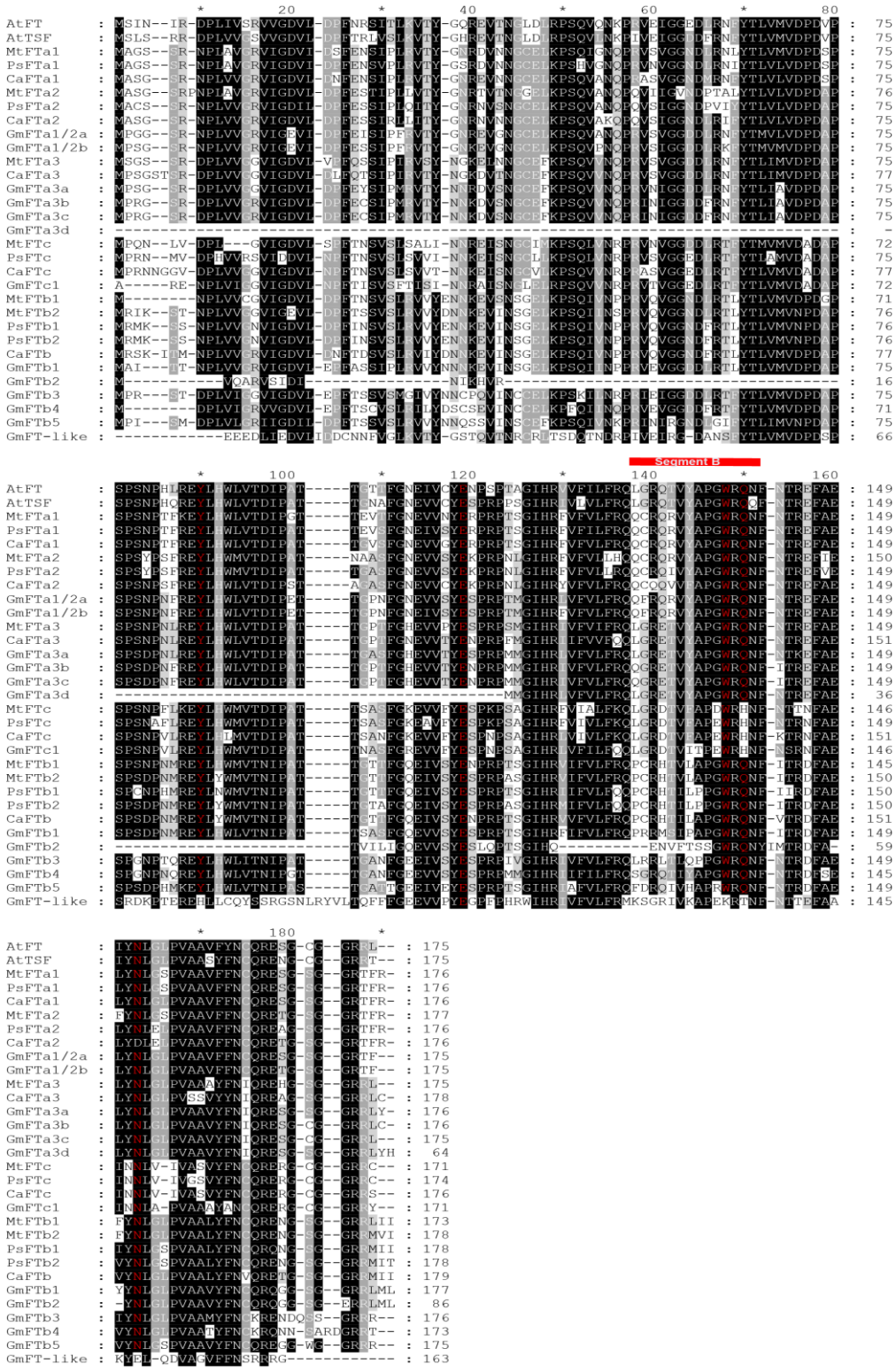


Figure 3.4 Multiple sequence alignment of FT proteins in five plant species. Protein sequences from accessions described in appendix 3.21 were aligned in Geneious 8 using MAFFT. Shades indicate conservation level: Black = 80%, grey = 60%, light grey = 40%. Important residues are highlighted in red (Ahn et al. 2006; Ho and Weigel 2014). Percentages of homology among proteins of different species can be found in appendices 3.23 to 3.25.

3.3.2 Distribution of Chickpea flowering-related genes in the genome

A high number of QTL analyses have been published on chickpea in the past two decades, many of which report QTLs for flowering time and other related traits. Consequent with the evolution of the chickpea marker pool, these studies employed many different types of markers such as simple sequence repeat (SSR), sequence tagged microsatellite markers (STMS), expressed sequence tags (EST), cleaved amplified polymorphic sequences (CAPS), Diversity Arrays Technology (DArT) or single nucleotide polymorphism (SNP). Although this evolution is positive in terms of mapping improvement, the lack of common markers makes the comparison between QTLs reported in different studies difficult.

Fortunately, the recent sequencing of chickpea genome enables the mapping of interesting sequences and place them physically in the chromosome. With this in mind, we reviewed the available literature and compiled all chickpea flowering-related QTLs and the markers defining them (summarized in Table 1.2). Then, we made an effort to position these markers on the chickpea genome to compare the results from different studies and find overlapping regions that could indicate the presence of major loci acting in different crosses.

Using this approach, we were able to position 86 markers linked or defining 36 QTL on the genome of cultivar CDC Frontier that was used as reference. Merging this information with the position of the 255 chickpea homolog flowering genes obtained by BLAST described in the previous section, we constructed a genome-wide map of the chickpea flowering candidate genes distributed across the 8 chickpea chromosomes (figure 3.5). The integration of both flowering-related genes and QTLs will be a useful tool that will help identifying potential candidate genes within the intervals of each QTL.

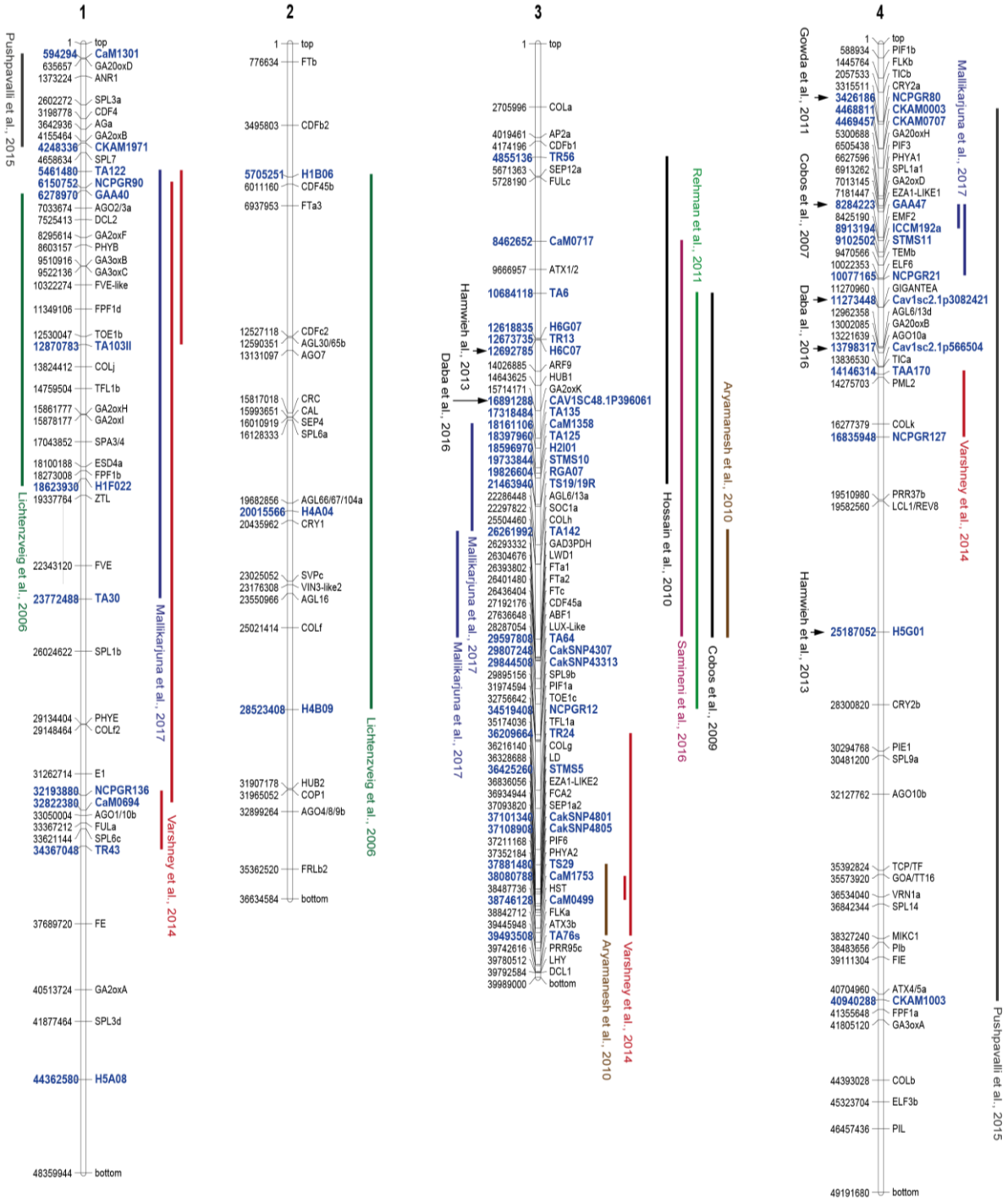


Figure 3.5 Graphical representation of the chickpea genome with the position of the 255 chickpea potential genes (in black) and some markers (blue) linked to flowering time QTLs (represented by bars or arrows). References are the same than those indicated in table 1.2. Bold numbers in the top indicate chromosome number, while numbers in the left of the bars correspond to physical position. First and last nucleotides of each chromosome are referred as "top" and "bottom", respectively. In each chromosome, bars of the same colour correspond to QTLs reported in the same study.

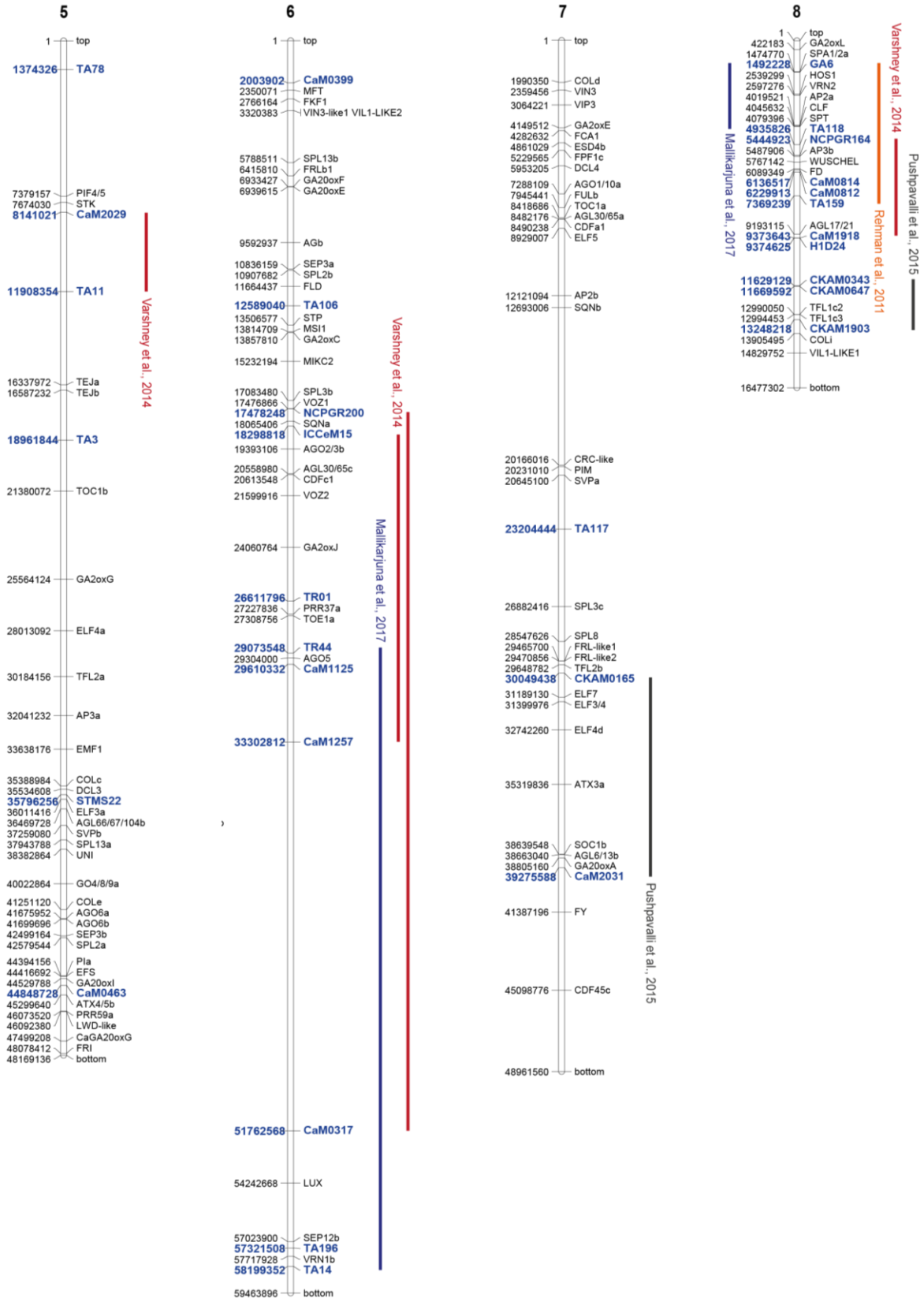


Figure 3.5 Continued

3.4 Discussion

Recent progress in development of genomic tools available for chickpea, and in particular the sequencing of the whole genome in 2013 (Jain et al. 2013; Varshney et al. 2013c), have opened up the possibility to rapidly identify chickpea genes homologous to genes with known function in other species, for any biological process. In the case of flowering, extensive studies have been performed in several model species identifying and characterizing many of the key genes. Among these species, the most relevant are *Arabidopsis thaliana* and the cereals *Triticum aestivum* (wheat), *Hordeum vulgare* (barley) and *Oryza sativa* (rice) (Greenup et al. 2009; Shrestha et al. 2014). Homologs of the genes described in these species have been shown to have conserved function and capacity to alter flowering time in other species. Hence, a comparative analysis looking for homologs of flowering-related genes is an efficient strategy to find potential genes involved in floral transition in species like chickpea where information on the molecular aspects of this process is lacking. Of all major model species, the most suitable for comparison with legumes is *A. thaliana*, due to its taxonomical proximity and also because it accumulates the majority of the knowledge on flowering time regulation. Moreover, the identity of the genes behind many legume flowering mutants correspond to orthologs of genes described in Arabidopsis (Liew et al. 2014b; Weller and Ortega 2015). In this study, 234 *A. thaliana* genes with potential roles in flowering time were used to identify a total of 255 homologs across chickpea genome. Most of the genes used as query were found to have close homologs in chickpea, and the results obtained from phylogenetic analyses were consistent with previous comparative studies in legumes. These general observations suggest that basic flowering pathways are likely to be well conserved between *A. thaliana* and legumes, and between chickpea and other legumes. Some of the genes included in the search have been shown to have important roles in flowering variation in the close legumes species pea and Medicago (see chapter 1, section 1.5.2), and it is very possible that the same scenario is also true for chickpea, so the information presented here will be valuable in future studies on flowering time in this important pulse crop.

A good understanding of the genetics regulating any agronomic trait is crucial for its effective manipulation. The identification of QTL controlling a trait and the isolation of the underlying genes has been established as one of the most effective approaches for dissection of complex

traits in crop plants (Pflieger et al. 2001; Zhu and Zhao 2007). Many strategies can be adopted to this end, but the high degree of conservation observed in the number and identity of flowering genes support the use of the candidate gene approach to find the molecular identity of the genes underlying QTLs reported for flowering time in chickpea. More than 50 of these QTLs can be count across all chickpea linkage groups (Table 1.2). Domesticated chickpea suffered several bottlenecks during its evolution, resulting in a very low genetic diversity (Abbo et al. 2003a), which is a limiting factor in the development of new molecular markers. Since QTL accuracy is very dependent on the population size and the number of markers, the low level of polymorphism between *C. arietinum* accessions can result in very wide QTL intervals in the case of intraspecific populations. Fortunately, the availability of the genome sequence (Jain et al. 2013; Varshney et al. 2013c) enables the physical mapping of the markers defining reported QTLs. This facilitates the analysis of potential genes within the interval delimiting each QTL to select the most plausible candidate genes in each case.

Candidates genes on Chromosome 3

Chickpea chromosome 3 accumulates the highest number of flowering time QTLs. More than 20 QTLs have been reported for this trait in at least 2 different regions of this chromosome (Table 1.2, Figure 3.5). The first region is close to the bottom (Varshney et al. 2014a; Aryamanesh et al. 2010), in the vicinity of several genes with links to light signalling and clock function, including a *PSEUDO REGULATOR RESPONSE* (*PRR*) homolog, a PHYTOCHROME-*INTERACTING FACTOR 6* (*PIF6*) gene, and a *PHYA* homolog (*CaPHYA2*). The second region is the most recurrent in the literature, reported in a wide variety of inter- and intraspecific crosses (Daba et al. 2016a; Hamwieh et al. 2013a; Mallikarjuna et al. 2017; Aryamanesh et al. 2010; Cobos et al. 2009; Samineni et al. 2016; Rehman et al. 2011; Hossain et al. 2010). By comparative analysis of the intervals in these studies, we can define this common region as delimited by markers TA6 and TA64. The aim of this chapter was to identify all the possible candidate genes within this region, in order to target them in future chapters attempting to uncover the molecular identity of the gene(s) responsible of the QTLs in this region. A total of eleven genes with potential functions related to flowering were found within the interval TA6-TA64. Two of these were MADS-box genes: *CaAGL6/13a* and *CaSOC1a*. Arabidopsis *AGL6* participates in floral meristem development (Dreni and Zhang 2016) and regulation of flowering time by repressing

FLC and enhancing *FT* expression (Yoo et al. 2011; Hsu et al. 2003), while *CaSOC1a* is one of two chickpea homologs of the floral integrator *SOC1*, a MADS-box gene with a central role in the control of flowering in Arabidopsis. Together with *FT*, *SOC1* acts as an integrator of the different flowering pathways to promote flowering by activation of floral meristem identity genes under the right conditions (Moon et al. 2003b; Lee et al. 2000).

Several genes with a potential connection to photoperiod responsiveness were also identified within the interval, including a *CYCLING DOF FACTOR* (*CDF*) gene, a *LUX*-like gene and a putative *LIGHT-REGULATED WD1* (*LWD1*) ortholog. *CDFs* are involved in the photoperiodic regulation of flowering time in several species (Kloosterman et al. 2013; Fornara et al. 2009; Ridge et al. 2016). *LUX-like* is a gene close in homology to *LUX ARRHYTHMO* (*LUX*), whose role and importance for flowering time will be discussed later in this section (when addressing the candidate genes found in LG6). In Arabidopsis, *LWD* genes encode a WD repeat-containing proteins that function in the light input pathway and are needed for the correct expression of the clock components *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*), *PRR9*, *PRR7* and *PRR5*. The *lwd1lwd2* double mutant has an early-flowering phenotype and altered length of the clock oscillation (Wu et al. 2016; Wu et al. 2008; Wang et al. 2011).

The list of candidate genes also includes a hormone related gene (*CaGA2ox*) and a gene involved in histone modification (*CaHUB1*). Gibberellins (GAs) can promote flowering in Arabidopsis through the activation of meristem identity genes (Blazquez et al. 1998), and mutants with low level of GA are unable to flower under non inductive conditions (Wilson et al. 1992). They have been implicated in the control of flowering in several other species, although there is no strong evidence supporting that this is the case in temperate legumes such as pea (Weller et al. 1997b). *GA2 oxidase* (*GA2ox*) belongs to one of the four clades of the *GA oxidase* gene family involved in GAs biosynthesis. Its function is to decrease the levels of active GAs by chemical inactivation (Hedden and Phillips 2000; Schomburg et al. 2003), so a mutation in this gene could potentially cause early flowering by GA accumulation (Huang et al. 1998), which would be more evident under short photoperiods.

Histone modification plays a crucial role in epigenetic control of transcription. The most common modifications include phosphorylation, methylation, acetylation and ubiquitination, and the balance between these marks dictates the epigenetic state of the chromatin. Since the switch

from vegetative to reproductive stage is one of the most important in any plant life cycle, genes controlling this process are also the targets of epigenetic regulation. In Arabidopsis, *HISTONE MONOUBIQUITINATION 1* (*HUB1*) and *HUB2* interact with other genes to upregulate the expression of *FLC* and *MAFs* through mono-ubiquitination of associated histones, therefore repressing the floral transition (Gu et al. 2009; Eckardt 2007).

A *COL* homolog to Arabidopsis *COL14* and *COL15* (*CaCOLh*) was found in a central position of the interval. As previously stated in the results section, *CO* and *COL* genes are of special relevance due to their role in a wide range of plant developmental processes, including flowering (Valverde 2011). However, opposite results have been obtained in legumes; while none of the *COL* genes tested in Medicago (including the ortholog *MtCOLh*) show correlation with *FT* expression or flowering induction (Wong et al. 2014), in soybean there is some evidence supporting an active role of these genes in flowering induction (Wu et al. 2014). In any case, the background of this family gene and the position of *CaCOLh* in the interval makes it a plausible candidate, so it needs to be targeted in future work to narrow the interval.

Finally, owing to their universal association with flowering time, the most plausible candidates for the QTL in the middle of chromosome 3 are the *FTa1*, *FTa2* and *FTc* cluster. In fact, most of the candidate genes discussed in this chapter influences flowering time by direct or indirect regulation of the *FT* gene.

Overexpression and complementation experiments found an ability to induce flowering in Arabidopsis background in almost all legume *FTs* tested, consistent with the observed conservation in amino acid sequence, and suggesting a conserved function (Ono et al. 2010; Laurie et al. 2011; Hecht et al. 2011). Physiological and mutant analyses also suggest a role of legume *FT* genes in vernalization and photoperiod responses, and different *FT* genes show differential tissue/photoperiod expression patterns. All together, these results indicate a diversification in the regulation and function of these genes during evolution of the different species.

Due to the close synteny between legumes, particularly high in those belonging to the same clade, the role of the 3 *FT* genes in the cluster (*FTa1*, *FTa2* and *FTc*) has been well studied in Medicago and pea. Of special relevance seems to be *FTa1*, which acts as a mobile floral signal

strongly correlated with flowering in both species (Laurie et al. 2011; Hecht et al. 2011). Another member of the *FTa* subclade (*FT2a*) has a similarly important role in soybean (Sun et al. 2011). In Medicago and pea at least, *FTc* expression is restricted to apex, but shows the highest potential in flowering promotion in complementation studies and could act redundantly with other *FT* genes through cross-regulation (Laurie et al. 2011; Hecht et al. 2011).

The chickpea PEBP family has been characterized in this study, finding a high degree of conservation between the genomic structure of the chickpea *FT* genes and their protein sequences with those from Medicago and pea, two of the most taxonomically close legumes. In view of the importance of the genes in the *FT* cluster reported in these species, it is reasonable to assume that their functions might also be conserved in chickpea, supporting the theory that these three genes are the most plausible candidates for the recurrent QTL reported between markers TA6 and TA64 (Weller and Ortega 2015).

Colocation of flowering QTLs and candidate genes on other chromosomes

Out of seven different QTLs reported, only five could be mapped to chickpea LG1, delimiting three different regions linked to flowering variation. Pushpavalli et al. (2015) described a QTL at the top of chromosome 1 with an interval that includes a *CDF* homolog, whose relevance has already been discussed above. A broad second region across the middle of chromosome 1 is defined by a QTL identified in three different studies (Lichtenzveig et al. 2006; Mallikarjuna et al. 2017; Varshney et al. 2014a), containing several potential candidate genes, among which *CaPHYB* and *CaCOLj* are perhaps the ones to highlight. Finally, Varshney et al. (2014a) found a third region containing a *CaFULa* gene. *FUL* homologs in other species are involved in flowering time control and meristem identity, so this gene represents a good candidate for this QTL (Berbel et al. 2012; Ferrandiz et al. 2000; Jaudal et al. 2015).

Only one of the three QTLs reported in chickpea linkage group 2 could be positioned in the genome (Lichtenzveig et al. 2006), delimiting a wide area of 22.8 Mbp. With such a large interval, the number of genes with potential to influence flowering is too high to provide any meaningful insight. In this scenario a first step narrowing the interval would be necessary prior to select candidates for further research. In any case, several genes across this region are worth mentioning, including two homologs of the above mentioned *CDFs*, a member of the chickpea

COL family (*CaCOLf*), the single chickpea *CRY1* homolog, a novel *FT* homolog (*CaFTa3*); a gene related to *SHORT VEGETATIVE PHASE* (*SVP*) which influences flowering response to ambient temperature (Jeong et al. 2007; Lee et al. 2007), and the floral meristem genes *CAULIFLOWER* (*CAL*) and *SEPALLATA4* (*SEP4*) (Zahn et al. 2005; Ditta et al. 2004; Purugganan and Suddith 1998; Kempin et al. 1995). Considering the known role of these genes and their close homologs in Arabidopsis and other species on flowering time regulation, they should be considered in future molecular analysis of the contribution of this region to flowering time variation.

Chromosome 4 is the second most prominent in terms of number of flowering QTL described; Mallikarjuna et al. (2017) reported two overlapping QTLs (GAA47-ICCM192a and NCPGR21-GAA47) in two different crosses. Since both involve CDC Frontier as maternal accession, the simplest initial hypothesis is that they represent the same underlying gene. GAA47 has also been associated with flowering regulation in the study of Cobos et al. (2007), and, in a third recent study, Daba et al. (2016a) also found several markers linked to flowering time in the neighbouring region. Among numerous flowering-related genes across this region, two genes initially stand out as the best candidates; *GIGANTEA* (*GI*), and one of the two chickpea *PHYA* homologs. Consistent with the role of PHYA as an important photoreceptor, *PHYA* genes have a major role in controlling the flowering response to photoperiod in several species, acting under long day (LD) conditions to inhibit flowering in the short-day legume soybean (Xu et al. 2013; Watanabe et al. 2009; Liu et al. 2008) and to promote flowering the long-day legume pea (Weller et al. 2004; Weller et al. 1997a). *GI* is a plant specific nuclear protein involved in diverse physiological processes that include a central role in the circadian clock and the flowering response to photoperiod in Arabidopsis. Loss-of-function *gi* mutants display altered circadian rhythms and cause delayed flowering time through several distinct mechanisms involving direct and indirect regulation of *CO* and *FT* genes (Mishra and Panigrahi 2015), whereas *GI*-overexpressing plants have an early, day-length insensitive flowering phenotype (Mizoguchi et al. 2005; Araki and Komeda 1993). A flowering-related role for *GI* orthologs has also been found in the legumes soybean and pea. The pea ortholog, *LATE1*, controls circadian rhythms and promotes flowering under LD, similar to Arabidopsis (Hecht et al. 2007b), whereas one of three *GI* orthologs in soybean, *E2*, inhibits flowering independent of photoperiod (Watanabe et al. 2011).

Another QTL in a distinct region lower on LG4 was described by Varshney et al. (2014a), delimited by markers TAA170 and NCPGR127. The only notable gene found within this region is a *COL* homolog (*CaCOLk*), homolog of Arabidopsis *COL6* and *COL16*. The relevance of *COL* genes has already been addressed in the case of *CaCOLh* (see candidate genes in LG3). Pushpavalli et al. (2015) also reported a QTL that we were able to assign to chromosome 4. However, the low number of markers used in their study resulted in a very wide interval for this QTL that occupies almost the whole chromosome (36.5 Mbp), offering no possibility to discriminate among any of the candidate genes identified, including those mentioned above.

Early flowering 1 (*Efl1*) is one of the four major flowering genes reported to date in chickpea. It was found initially by Kumar and Van Rheenen (2000) in the chickpea genotype ICCV2. Since then, it has been linked to chromosome 5 in several studies (Pushpavalli et al. 2015; Vadez et al. 2012; Jamalabadi et al. 2013; Cho et al. 2002). Recently, Ridge et al. (2017) showed that this gene is likely to be a chickpea ortholog of Arabidopsis *ELF3*. Besides those pointing to *Efl1*, one more QTL has been reported by Varshney et al. (2014b) between markers CaM2029 and TA11. None of the 255 chickpea genes retrieved in the present study was located within this interval, so in this particular case a deeper analysis of the genes annotated between these two markers may be required in future.

Three different studies have reported QTLs for flowering control in chromosome 6. Varshney et al. (2014a) found a recurrent QTL across different seasons and locations, in a region containing several candidate genes, among which a *PRR*, a *CDF* and *a TARGET OF EAT1* (*TOE1*) homologs seem the most promising due to their established roles in photoperiod response and clock function (Nakamichi et al. 2010; Zhang et al. 2015; Hayama et al. 2017). Mallikarjuna et al. (2017) reported a QTL in a region closer to the bottom of the chromosome between markers TR44 and TA14. Four genes potentially related to flowering time control were mapped within this interval, but three of them (*AGO5*, *VRN1b* and *SEP12b*) were close to the limits of the interval and therefor unlikely candidates. The best candidate, due to its central position in the QTL is the chickpea ortholog of *LUX*, a Myb-domain transcription factor necessary for the correct function of the circadian clock and acting to delay flowering under non-inductive LD (Helfer et al. 2011; Hazen et al. 2005). The role of *LUX* orthologs appears well conserved in other species (Huang and Nusinow 2016; Campoli et al. 2013), including the close legume

species pea where mutations in the ortholog *SN* confer early, photoperiod-insensitive flowering (Liew et al. 2014a).

Linkage group 7 seems to be the least relevant for flowering induction in terms of reported QTLs. Only two QTL have been mapped to this LG, and only one of these, described by Pushpavalli et al. (2015), could be physically located, between markers CaM2031 and CKAM0165. Among other potentially relevant genes, chickpea homologs of Arabidopsis flowering genes *ELF3*, *ELF4* and *SOC1b* were also found in this region. *ELF3* and *ELF4*, together with *LUX*, form the so called evening complex of the circadian clock, which coordinates environmental and endogenous signals and thus participates in the regulation of several processes (Ezer et al. 2017; Doyle et al. 2002; Zagotta et al. 1996; Liew et al. 2009; Weller et al. 2012; Rubenach et al. 2017; Melzer et al. 2008).

Several QTLs have been linked to chromosome 8. The comparison the position of their intervals point to at least two possible different candidate genes; Pushpavalli et al. (2015) described a narrow QTL close to the bottom of chromosome 8 between markers CKAM1903-CKAM0343. Only one candidate gene (*CaTFL1c*) was found within this interval. *TFL1* genes are flowering repressors, and *tfl1* mutants show early flowering in different species (Varkonyi-Gasic et al. 2013). Moreover, Hecht et al. (2011) proved that the flowering promotion of FT proteins in pea involve down-regulation of *TFL1* homologs, so this gene is a strong candidate for the QTL in this region and should be targeted in future works. Mallikarjuna et al. (2017) and Rehman et al. (2011) described two QTL that co-locate perfectly, therefore this two studies are likely reporting the same gene. It could also be the same QTL described in Varshney et al. (2014a), although in this case the interval is slightly displaced towards the central part of the chromosome and could be signalling to a different, third locus.

Concluding comments

The aim of this chapter was to investigate the co-location of reported flowering QTLs with 255 genes, chickpea homologs of a selection of genes involved in floral transition in Arabidopsis. It has provided a preliminary analysis of the most suitable candidates for the different QTLs, focusing on QTLs in an 18.9 Mbp region of chickpea chromosome 3 reported from several different inter- and intra-specific crosses. We confirmed the presence within this region of a

cluster of 3 *FT* genes, which had been proposed as the most plausible candidate by Weller and Ortega (2015), but also identified 10 other genes that could potentially explain the flowering differences observed in the different environments and populations. In future studies, the region needs to be saturated with new markers, and as many as possible of all candidate genes discussed here need to be targeted in different crosses, in order to narrow the QTL interval. This will allow to discard some of the potential genes (all but one in an ideal scenario) and would help to answer the questions of 1) whether this region associated with flowering time in different chickpea populations represent the same gene or there is more than one locus acting and 2) the identity of the responsible gene(s).

Chapter 4. QTL analysis of flowering time in chickpea

4.1 Introduction

Many QTL have been reported to influence flowering time in chickpea, located on all eight chickpea chromosomes (summarized in Chapter 1, Table 1.2**)**. In chapter 3 of this thesis, we presented a broad genomic overview of the relative position of potential flowering-related genes and QTL by analysis of the genomic location of flanking markers, and identified several homologs of Arabidopsis flowering-related genes as potential QTL candidates, focusing in particular on a region of LG3, that has been associated with regulation of flowering in 14 different QTL studies and therefore seems especially important for this trait. The presence of a *FTa*/c cluster within the region and the fact that QTLs controlling flowering are located in the syntenic region in other legume species led to the proposition of this cluster as the most plausible candidates (Weller and Ortega 2015).

The main objective of this chapter is the evaluation of the *FTa*/*c* cluster as candidates controlling flowering in chickpea using two different approaches. First, we will test the association of the *FT* genes with flowering time in an interspecific population with a reported flowering QTL in the above-mentioned region [RIP12, Cobos et al. (2009)]. Then, we will attempt to delimit the QTL with new gene-based markers, targeting where possible those potential candidates obtained in Chapter 3. We will examine the importance and nature of this locus by analysing the genetic control of flowering in three more populations, one interspecific and two intraspecific. To this end, we will use common markers in the four populations as anchors that will enable us to analyse the co-location of the obtained QTL.

In parallel, we will analyse the expression of the *FT* genes in the parental lines of the crosses used to demonstrate the association. The PEBP family have been molecular and physiologically characterized in species taxonomically close to chickpea such as *Medicago truncatula* and *Pisum sativum*. In these legumes, the role of the FT proteins as integrators of environmental cues and promotion of flowering have been demonstrated (Hecht et al. 2011; Laurie et al. 2011; Foucher et al. 2003). Due to their proximity in the legume clade, we expect this floral promoter function to be conserved also in chickpea, so any difference in flowering time could be linked to differential expression of the *FT* genes.

In view of the importance of this gene family, a secondary objective of this chapter is to characterize the developmental expression profile of the chickpea *FT* genes, as well as other genes associated with floral induction, in order to investigate the extent to which the regulatory patterns in chickpea are conserved with those in other legumes.

4.2 Materials and methods

Plant material

Four recombinant inbred populations derived from both interspecific and intraspecific crosses were used in this study. Basic information about the crosses is given in Table 4.1.

Table 4.1 Summary of the populations used in this study

Name	Cross	RILs (N)	Developer
Interspecific populations			
RIP12	ICCL81001 x Cr5-9	88	Department of genetic, University of Cordoba (UCO)
CRIL2	ICC4958 x PI489777	124	USDA[a] and Washington State University (USDA-WSU)
Intraspecific populations			
RIP5	WR315 x ILC3279	102	Department of genetic, University of Cordoba (UCO)
RIP8	ILC3279 x WR315	113	

a. United States Department of Agriculture

RIP12 (Recombinant Inbred Population 12) is an interspecific population consisting of 88 $F_{6:7}$ RILs (Recombinant Inbred Lines) derived from across between the *kabuli* cultivar ICCL81001 and *C. reticulatum* (Cr5-9). This population was previously described by Cobos et al. (2009), where further information can be found.

The CRIL2 population is a second intraspecific population comprised of 128 $F_{6:7}$ RILs from the interspecific cross between *C. arietinum* accession ICC4958 and *C. reticulatum* (PI489777). It was developed by the United States Department of Agriculture, Agricultural Research Service and Washington State University (Pullman, Washington, USA), and is now generally considered as a reference population, in part because of extensive marker sets and the availability of genome sequences for both parents. ICC4958 is a desi chickpea type from India maintained by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) with an erect growth habit and an early flowering phenotype, while the wild parent PI489777 is an accession from Turkey with a prostrate growth habit and late flowering phenotype typical of wild chickpea.

Two further $F_{6:8}$RIL populations derived from the intraspecific cross between WR315 and ILC3279 (RIP5, composed by 102 RILs) and its reciprocal cross (RIP8, 113 RILs) were developed and provided by researchers of the Department of genetic at the University of

Cordoba (Table 4.1). ILC3279 is a late-flowering *kabuli* type from the former Soviet Union maintained by the International Centre for Agricultural Research in the Dry Areas (ICARDA), while WR315 is an early flowering *desi* landrace from central India maintained by ICRISAT (Ali et al. 2015; Iruela et al. 2007).

Growing conditions and phenotypic evaluation

RIP 12

RIP12 was sown in the field at IFAPA Centro Alameda del Obispo (Cordoba, Spain) over four different seasons (2001, 2004, 2008 and 2014) on 3rd, 13th and 11th of March, respectively. In all four years and environments, flowering time (DTF, also considered as flowering initiation) was calculated as the number of days from sowing until 50% of the plants in a plot had open flowers. Growing conditions can be found in Cobos et al. (2009). In addition, two other traits were recorded from the 2001 trial and are included in the present QTL analysis study. The end of flowering (End) was measured as the number of days from sowing until 50% of the plants in a plot had developed green pods, and the days to maturity (Maturity) was recorded as the time from sowing until 50% of the plants carried dry pods containing mature seeds.

CRIL2

For each RIL and the parental lines used as controls, four plants (two pots with a density of 2 plants/pot) were sown and grown under either long day (LD) or short day (SD) photoperiod, as described in chapter 2, in the School of Biological Sciences phytotron at the University of Tasmania (Hobart, Tasmania), between December 2015 and April 2016. Plants in each treatment were arranged in a completely randomized experimental design.

Flowering time was recorded as the number of days from seedling emergence to opening of the first flower (DTF) on each individual plant for each RIL under both photoperiod treatments. The mean of all plants in each RIL was calculated and used for analysis.

RIP5 and RIP8

RIP5 was sown in 2003 at two different sites: at IFAPA Centro Alameda del Obispo (Cordoba, Spain) (sown on March 14th) and at Mengibar (Cordoba, Spain, sown on March 11th). In each

assay, RILs were distributed randomly into four blocks (20 lines per block). Four check lines were included in each block following a Latin square design to verify environmental homogeneity. The plot unit was three rows, 4 m long, with 0.5 m between rows and a density of 20 plants m^{-2}. RIP 8 was sown the 4th of February 2003 using the same experimental design at IFAPA site with two replications of the trial. For both populations, days from sowing to 50% flower was recorded (DTF) and the data obtained from each of the two trials of RIP8 were analysed separately.

Molecular Markers

All four populations were grown in Hobart, and DNA was extracted following the protocol described in the General Materials and Methods (Section2.3.1). For the purposes of this study, polymorphisms in genes across chickpea linkage group 3 were identified in each of the four populations (Table 4.2) either by sequencing of target genes in the parental accessions or from previous reports (Saxena et al. 2014). HRM markers were designed to target SNPs or small InDels (Table 4.2). Sequencing and scoring of the markers were performed as described in Sections 2.7 and 2.8.

The genetic map of the RIP12, RIP5 and RIP8 populations have been described in Cobos *et al.* (2009), Ali et al. (2015) and Iruela et al. (2007), respectively. They were revised using the same marker sets but with new mapping conditions (described in next section) and the addition of new specific HRM markers described in table 4.2.

For CRIL2, a subset of 210 markers covering 9 linkage groups was selected from a denser map based on 2956 markers, developed and kindly provided by Prof Doug Cook and Dr. Varma Penmetsa in the Department of Plant Pathology of University of California, Davis (Davis, USA). Markers were first filtered to minimize data loss and then the selection was made, mainly based on chromosomal distance; in general, one marker was chosen every 5 cM. The final set (described in Appendix 4.1) had an average distance between adjacent markers of 2.5 cM. The 14 specific HRM markers developed for this study were also scored in this population for integration in the map (Table 4.2).

Table 4.2 Details of target genes and new HRM markers developed for analysis of the LG3 region in the different chickpea populations.

Marker	Population	Fw Sequence	Rv Sequence	GeneID
LHY	CRIL2/RIP12	AAACCAACTAAGCATACCCT	TGAGCATCACTCATTCACCA	101500635
	RIPs8/5	CGTCACACTTGTAATCTTCATTCC	AGTTTTCCCCCTTTAATAATGTGG	
CP450	CRIL2/RIP12	CACAAAATAGAGAACAATGACAGC	ACTTTTCCCTTTGCATGTAGG	101502624
	RIPs8/5	AGAGTTGTATAGTTGTTAAGGATG	GTGTGTGTGTTTATCAATTTAAGC	
CDF2d	CRIL2/RIP12	TGGTTCCAATTAAGTTTCAAGTG	AGTCAAGTGTTTGGTAAGAGTTG	101500722
	RIPs8/5	AGTCGATGCTTAATCTTCAACAGC	AGATCTGCATAAAGATGGTTCC	
WUS11	CRIL2/RIP12	CAGCCTGGTAATTAACTGCATC	ATGATTTTGAGCAATTATTCTGTG	101503157
	RIPs8/5	GCATAACCTAGAGTGATCGAGC	CTACTCTGACTTAATGGGTTCG	
FTa1	CRIL2/RIP12	AATCATCCCCAAGAGATCAA	TGCACAGTCATTGTGTTTCG	101497376
	RIPs8/5	ACTGTTCTGCACACAGTGGCTACC	AAAGAGATCTAACACATTTTGC	
			GAGGAACACGACGCCTCC	
SUVH4[1]	CRIL2/RIP12	TTCGCCGTCACTACCTCG	GGAGATTAAGCTTCGGAGG	101508428
COLh	CRIL2/RIP12	TGCTACATCATTCACCTAGTAACA	TGCCATGATATAGGAAGTCTTAGTT	101504031
SOC1a	CRIL2/RIP12	AAACAAAGAAAAACGAATGTGTCC	CGACATATAATTCATTGTGGACCG	101510775
SCARECROW[2]	CRIL2/RIP12	GAGACATGTTGTTGAACAGC	CTTGATGGTCCTCTAACAGC	101513767
RAPTOR1[2]	CRIL2/RIP12	CCCAATGCCATCCAAATCGG	CACACAACAACAATACCTAGGG	101514864
ARF9[2]	CRIL2/RIP12	GCAATATGGTGAGAAGAATTTTC	TGAGATAGGCAATTTAGTCCCTG	101491204
ABI5	CRIL2/RIP12	TTAGCACAAGAACGAGCGAGC	GTGAACGAAAAGGTGTTATGAAGG	101490892
GATA9[1]	CRIL2/RIP12	GAGGAATATGTCTTCTTCCATTCC	TGGAAAGAGTAATTTTCCCCCTA	101503040
NAC100[1]	CRIL2/RIP12	CAGGTCTTAGCAATGACACG	GCCCTATTTCTTCCCATGTC	101500623
CDF3b	RIPs8/5	AACAACCGAAGAAAAATAGG	GAATTTGTATAATGTTTATCTTCG	101499964
COLg	RIPs8/5	AGAGACTCTGAAGGTGTCCC	CAGTGGCTCGGAGAAAGTGG	101499146
AP2-like	RIPs8/5	AACAAACATCGTCACATCACC	CCTTGTGCTATTTAGTGTTCTGC	101502947
LOB189	RIPs8/5	ACAAAATCAATACAAGCAAACC	TGCAACCGTTAGTTTGTTTGG	101508422
PRSP	RIPs8/5	GATACATTTTCGCTCAAACTATG	TGCGTTGAAAAGTGTTTTATTAGC	101491522
PRT6	RIPs8/5	AAATTTTCATTCTCCTAAGACAGT	ACGGTCCAACCCAACGTATA	101506928
WRKY	RIPs8/5	TTCTGAGAGCACCGTGATGG	AGCATCTCCAACTGTAATTAAATG	101511519

(1) Primers designed by Saxena et al (2014)
(2) Primers designed for this study based on SNPs described by Saxena et al (2014)

Genetic mapping and QTL analysis

Linkage analyses were performed using JoinMap v4.0 (Van Ooijen 2006) using similar mapping strategies for all populations. The independence logarithm of odds (LOD) with a minimum value of 3.0 was used to define the linkage groups. The regression algorithm with JoinMap default options (recombination frequency <0.40 and LOD >1, goodness-of-fit jump threshold 5.0 and ripple 1) was used for mapping, using the Kosambi function for estimation of distances (Kosambi 1943). The initial maps obtained were reviewed and problematic markers were removed where necessary based on the Chi square contribution to the goodness-of-fit (with a maximum threshold value of 1.0) and the level of marker segregation distortion compared to surrounding markers.

The identity of the linkage groups obtained from this analysis was established according to the presence of common markers with the linkage groups described in the chickpea consensus

genetic map (Millan et al. 2010), and that correspond with the eight chromosomes of the chickpea genome. In case where markers from individual populations were not present in the consensus map, the Cool Season Food Legume Database (https://www.coolseasonfoodlegume.org) was used to find other published chickpea maps containing markers that are linked to both the markers in the unknown group of this study and those in the consensus map, allowing us to infer the corresponding chromosome.

QTL analysis was performed using MapQTL 6 software, according to the following protocol. First, interval mapping was carried out to detect putative QTL associated with the trait variation. For each putative QTL, the marker with the highest LOD score together with the two adjacent markers in each direction of the map were then used for Automatic Cofactor Selection (ACS). The marker designated by ACS was used as a cofactor for Multiple QTL Mapping (MQM) analysis. This second round of QTL analysis reduces the residual variance attributable to a previously detected QTL and increases the power of the QTL analysis to find other segregating QTLs for each studied trait. The MQM function was employed reiteratively with each new cofactor selection until all QTLs for a specific trait were determined. The threshold LOD necessary for a 0.995 QTL significance was estimated for each trait and linkage group by Permutation Test with 1000 permutations at a significance level of $p<0.05$. The LOD score peaks were used as estimators of QTL position in the map. The amount of variation explained by the markers was determined using the coefficient of determination (R^2) value and expressed as percent phenotypic variance explained (PEV%). Any QTL with a PEV higher than 10% was considered as major.

RNA extraction and qPCR.

The six parental lines of the four populations were grown in the phytotron at University of Tasmania under short and long day conditions. To analyse the developmental expression pattern of flowering-related genes, plant material from both photoperiods was collected every two weeks from emergence until plants were 12 weeks old in the case of ICC4958 and PI489777 (CRIL2 parental lines), or on a weekly basis from emergence during 4 weeks for WR315, ILC3279, ICCL81001 and Cr5-9 genotypes (RIPs 5/8/12 parents). Tissue was collected separately from dissected apical buds and uppermost fully expanded leaflets for analysis in these two tissues.

Each sample consisted of pooled material from 2 plants, and two biological replicates were performed. Each data point represents the mean ± s.e.

RNA extraction and cDNA synthesis were performed as described in Chapter 2, section 2.3.2. The expression of all chickpea *FT* genes as well as the floral indicators *PIM* and *UNI*, and other flowering related genes were measured as described in chapter 2.4.3. Primer sequences used in this chapter are indicated in Table 4.3. The primers used to amplify the housekeeping gene (B-actin) were obtained from previous expression studies on chickpea, while primers for all other genes analysed were developed specifically for this study.

Table 4.3 Sequence and product size of primers used to measure the expression of flowering related genes in chickpea

Gene	Fw primer	Rv primer	Product size (bp)		
			Genomic	cDNA	GeneID
ACTIN	ATTGTCTTGAGTGGTGGTTCT	TTCCTCTCTGGTGGTGCTAC	(Verma et al. 2013)		
PIM	GAACTTCAGAGTCTGGAACAGC	CATTGTGCCTGTTGTTGAGC	472	188	101488241
UNI	TGCAACGCGTAACAGTGAACG	ACGAACAATGCCGTGAGTTCTTG	660	169	101503680
FTa1	TTGCCAATCAACCCAGAGCG	AGTGGGGTTACTTGGGCTAGG	236	101	101497376
FTa2	GTTCTGACGGTGGTTCTCTC	CGGAGGTTCACAAAAGAAGG	297	183	101496618
FTc	TGTTGGTGGTGAAGATCTAAGG	ATTCCTGCTGAAGGATTCG	6616	186	101508200
FTb	GGTGAGCTCAAACCCTCCCA	TCCCTCATATTTGGGTCACTAGG	618	131	101505276
FTa3	CCATCCCGGAGCATACACAGTC	TGCACCAAGCCCTAGCAATCC	1460	196	101515383
E1	AACGACAACAACAAGGGATCGG	TGCAGCCAACAAGAGTCTGC	120	120	101497661
TFL1b	TATACCGGGCACAACAGATG	GGAGGTCCAAGGTCATTGTC	584	195	101508699
TFL1c	ACGTTCCTGGCCCAAGTGAT	TGGATCCCTATGTTAGGCTTTGGT	1134	136	101491943

4.3 Results

4.3.1 New markers and revised genetic maps

Across the four populations, 21 additional markers were designed on genes with known position according to the reference genome of the *kabuli* cultivar CDC Frontier [(Varshney et al. 2013c), BioProject PRJNA190909]. With the information about the position of these new markers and those flowering-related markers and genes obtained in chapter 3, we built a map of chromosome 3 (Fig 4.1), which provides a useful as a reference against which the quality of genetic maps developed in this study can be assessed and also for specifying the genomic intervals containing QTLs. Also, this will help determine the possible co-location of QTLs between populations or different traits, both within this present study and also among previously published reports.

There were two distinct aims of this exercise. The first was to narrow the interval for the flowering time QTL previously identified in RIP12 between markers TA6 and TA64 (Cobos et al., 2009). Therefore, the majority of the markers designed here are targeting genes within this region, either because of their potential as candidates (as they had functions related to flowering in other systems) or the region between them to reduce the inter-marker distance (Fig 4.1). In chapter 3 we suggested the *FTa-c* cluster as the strongest candidates for this QTL, so to test this hypothesis we included a marker targeting these genes in the four populations. Second, we used common markers across the four populations as anchor markers allowing us to compare the identity of the obtained QTLs. The four linkage maps corresponding to the four populations are presented in appendices 4.2 to 4.5.

RIP12

The linkage map derived from the reanalysis of RIP12 comprised a total of 155 markers grouped in 11 linkage groups covering a total distance of 178.7 cM, with an average inter-maker distance of 1.27 cM (Appendix 4.2). This same marker set was previously employed by Cobos *et al.* (2009), and the linkage groups were largely consistent with those reported in that study, with two exceptions. First, LG1 of Cobos et al. (2009) was not distinguished using the new mapping conditions. Second, markers TS45, TA25 and TA3 were included in the unlinked group 2 from Cobos *et al.* (2009). However, these three markers clearly correspond to chickpea LG8 (Millan *et al.,* 2010), so the linkage group was renamed LG8. A major QTL for flowering time was found

in this population between markers TA6 and TA64 on LG3 (Cobos *et al.,* 2009). Fourteen HRM markers were included to narrow this QTL (Table 4.2), and all mapped as expected according to the physical map (Fig 4.1). The HRM markers designed in the chickpea homolog of the circadian clock-related gene LHY was placed in the LG3B of Cobos *et al.* (2009). Since the physical position of this gene is known to be in the bottom of the chromosome 3 (Fig 4.1), LG3B and LG3A in the former map were renamed here as LG3a and LG3b respectively, to reflect the physical position of the markers in chromosome 3.

CRIL2

CRIL2 is considered as the reference mapping population for interspecific crosses between *C. arietinum* and *C. reticulatum*, with several high-density map already existing and reviewed by Sharma et al. (2013). To construct a lower-density map for the purposes of the present QTL analysis, 210 evenly-spaced markers were selected from the high-density and supplemented by markers for 14 of the LG3 candidates listed in Table 4.2. Two of these markers (ARF9 and NAC100) were removed for map refinement, so the final map has 222 markers on nine linkage groups, covering 540.3 cM, with an overall inter-marker distance of 2.56 cM (Appendix 4.3). The number of markers per linkage group ranged from 11 (LG9) to 39 (LG3), and their order in the new map perfectly match the original one (Appendix 4.1), which is indicative of the accuracy of the mapping method used in the present study. Similarly, the relative lengths of the linkage groups, ranging from 10.3 cM (LG9) to 83.9 (LG4), are also comparable in both maps.

RIP5/RIP8

While the new maps described for the RIP12 and CRIL2 interspecific populations were essentially derived from previously reported maps, the maps developed for the intraspecific populations RIP5 and RIP8 have never been published. The genetic map obtained for RIP5 is the least dense of the four, comprising 64 markers distributed across 7 linkage groups, with a total coverage of 87.3 cM and an average inter-marker distance of 1.53 cM. Due to the low number of markers used, linkage groups 1, 5 and 7 from the chickpea reference map were missing. Linkage groups 2 and 3 had the highest number of markers (24 and 21, respectively), whereas LG6 had the lowest, and was represented by only two (Appendix 4.4). For RIP8, 95 markers were grouped in 10 linkage groups covering 142.7 cM and the average spacing between consecutive markers

was 1.76 cM. All ten groups could be assigned to chromosomes according to the reference map, even though chromosomes 2, 3, 4 and 6 were each represented by two unlinked groups (Appendix 4.5). The highest number of markers mapped to LG4, 2 and 3 (21+9, 25+4 and 19+3, respectively) while LG5 and LG8 had only 4 markers each.

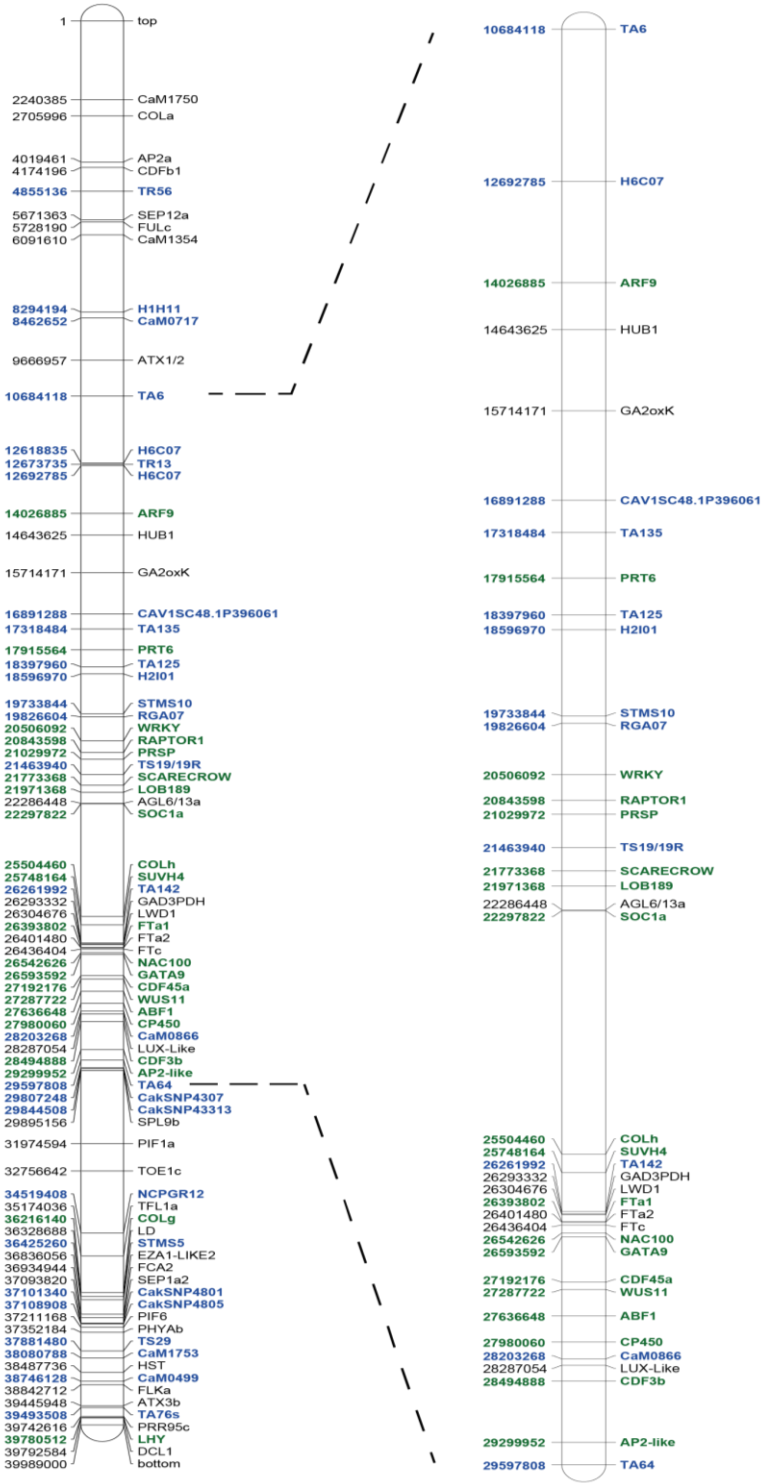


Figure 4.1 Representation of the entire chickpea chromosome 3 (left) and a magnified view of the region between markers TA6 and TA64 (right). Numbers on the left of each diagram indicate physical position with the first and last nucleotides of the chromosome referred as "top" and "bottom", respectively. Markers used in previous studies to position QTL (Table 1.2) are shown in blue, new markers designed in this study targeting flowering-related and other genes are shown in green, and several other genes of potential relevance to flowering are shown in black.

4.3.2 Phenotypic assessment

Figure 4.2 summarizes the values for flowering-related traits that were recorded from the four RIL populations and parent lines in different environments and seasons.

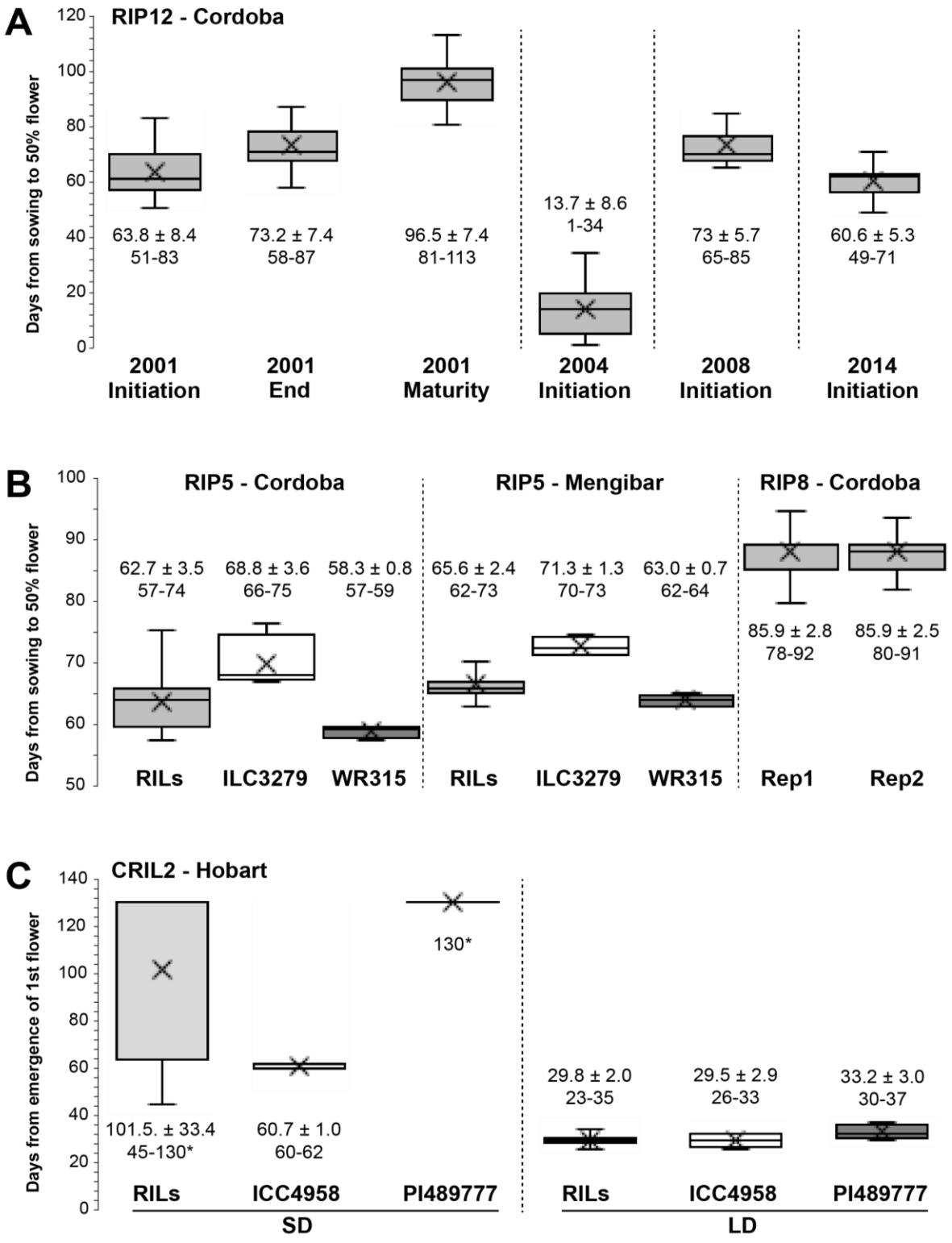


Figure 4.2 Phenotypic variation for flowering time across the four populations and environments. Numbers represent mean (± SD) and range. The asterisk for CRIL2 under SD indicates that plants did not flower by the end of the experiment and were assigned a nominal minimum flowering date of 130.

RIP12

Flowering time in RIP12 was evaluated in four different seasons. In 2004 the first RIL to flower was scored with a value of 1 and used as reference for the rest of the population. The absence of absolute values with this different scoring system makes it difficult to make inter-seasonal comparisons, but the differences between the earliest and latest lines (32 days) were similar to those obtained in 2001 (34 days), and substantially larger than in 2008 and 2014, where the range in flowering time was only 20 and 22 days respectively. Significant differences in flowering time were also found among years, with RIP12 flowering, on average, 9 and 12 days earlier in2001 and 2014 compared to 2008 ($t(78) = -10.695$ and $t(82) = 22.333$, $p < .001$). Differences between 2001 and 2014 were small (3.2 days) but significant ($t(82) = 3.291$, $p < 0.5$). Since sowing dates were almost identical across years, this variation reflects a strong influence of the environment on flowering time. A strong, positive correlation was found between flowering initiation, end of flowering and maturity during 2001 (Table 4.4), indicating that the lines with an early flowering were also early to mature and showing that flowering time could be used in this case as an accurate estimator of cropping duration.

Table 4.4 Spearman's rho correlation coefficient between days to 50% flowering (Initiation), days to end of flowering (End) and days to Maturity in RIP12 during 2001.

		End	Initiation	Maturity
End	Correlation Coefficient	1.000	.764**	.840**
	Sig. (2-tailed)	.	.000	.000
	N	84	84	82
Initiation	Correlation Coefficient	.764**	1.000	.728**
	Sig. (2-tailed)	.000	.	.000
	N	84	84	82
Maturity	Correlation Coefficient	.840**	.728**	1.000
	Sig. (2-tailed)	.000	.000	.
	N	82	82	82

**. Correlation is significant at the 0.01 level (2-tailed).

CRIL2

The difference in mean DTF between the CRIL2 parent lines grown under LD conditions was small (3.7 days), and not significant. The mean value for the RILs was intermediate between the two parent lines, and the distribution was strongly skewed towards the domesticated parent

ICC4958, which may explain why the distribution, despite appearing similar to a normal distribution (Fig 4.3 B), failed normality test ($W(125) = 0.963$; $p < .05$). The range displayed by the RILs was substantially wider than the parents, with 12 days difference between the minimum and maximum value for DTF. Under SD conditions, the flowering responses between the two parents were significantly different ($t(6) = -144.6$; $p < .001$): while ICC4958 flowered at 60.7 days, *C. reticulatum* failed to flower after 130 days. This distinct difference was reflected in the RILs, which showed a clear bimodal distribution, with either a flowering or a non-flowering phenotype (Fig 4.3 A). Paired t-test revealed significant differences in CRIL2 flowering time when comparing photoperiods, in both parents and the RILs ($t(3) = 17.8$, $t(3) = 64.8$ and $t(122) = -24.179$ for ICC4958, PI489777 and RILs, respectively; $p < .001$ in all cases). As expected in a LD species, all lines flowered considerably later under SD photoperiod. Flowering time in the RILs showed a weak but highly significant correlation (Spearman's rho; $rs[125] = .319$, $p < .001$), evidencing that part of variation is independent of photoperiod. Transgressive segregation, mostly towards earliness, was found in both conditions, and was more evident under SD, where some of the RILs flowered 15 days earlier than the early parent.

RIP5/RIP8

The RIP5 population was grown in two sites with minimal differences (3 days) in sowing time, and evaluated for days to 50% flower during 2003 season. In both locations, statistically significant differences were found between flowering time of the parental lines ($t(6) = -5.227$ and $t(6) = -9.661$ in Cordoba and Mengibar, respectively; $p < .01$ in both cases). Paired t-test revealed that WR315 ($t(3) = -6.3$; $p = 0.008$) and RILs ($t(3) = -6.3$; $p = 0.008$) flowered slightly earlier in Cordoba compared to Mengibar (4.7 in the case of WR315 and 3 days difference for the RILs), whereas ILC3279 does not show any difference between sites. In both locations, mean and range of the flowering time in the RILs had values intermediate between those of the parents WR315 and ILC3279, although the distribution was wider in Cordoba (Fig 4.2 B). As described for CRIL2 above, the flowering of the RIP5 RILs was skewed toward earliness (Fig 4.3 C and D).

Although grown in the same location (Cordoba) and season (2003), the RIP8 population was sown 38 days earlier than RIP5, on 4th of February. Not surprisingly, mean flowering time (quantified as days to 50% flowering) in RIP8 was considerably later (23 days) than RIP5 (Fig 4.2 B). However, in other respects, the two repetitions of the trial showed similar mean values

and ranges, indicating the relative homogeneity of the environmental influences and general quality of the two trials. In view of this homogeneity, the data from both repetitions were pooled together for graphical representation in Figure 4.3 (E).

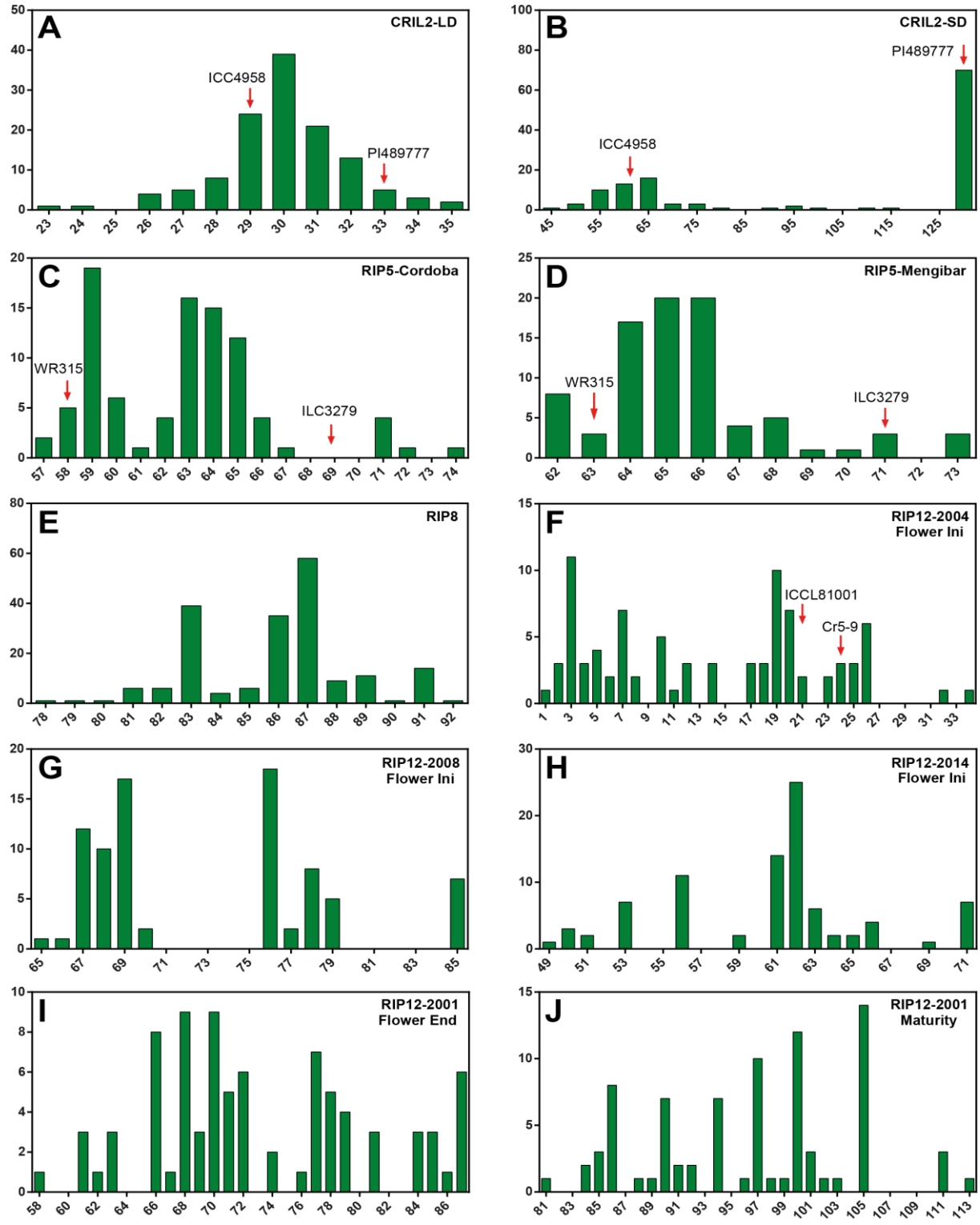


Figure 4.3 Frequency distribution for flowering time in four different RIL populations in different seasons and conditions. Vertical axis indicates frequency and horizontal axis flowering time measured as follows: Days to first open flower in CRIL2 grown under LD (**A**) and SD (**B**) conditions, Days to 50% flower in RIP5 at Cordoba (**C**) and Mengibar (**D**), in RIP8 (**E**) and in RIP12 during 2004 (**F**), 2008 (**G**) and 2014 (**H**). Days to end of flowering (**I**) and days to maturity (**J**) in RIP12 during 2001 season.

4.3.3 Quantitative trait locus (QTL) analysis

QTL analysis was performed for each relevant trait on each population in each separate trial or experimental condition. A total of 16 significant QTLs were identified, distributed in three regions of chromosome 3 and one of chromosome 4. 14 QTLs were obtained for flowering initiation, one for end of flowering and another one for maturity. A region of chromosome 3 containing the markers targeting *COLh* and *FT* genes seems especially relevant for chickpea phenology, harbouring a cluster of 10 QTLs for flowering initiation, end and maturity. A summary of relevant parameters is shown in Table 4.5.

Table 4.5 Summary of the QTLs found for each population and trait analysed in this study indicating the conditions, years and places, as well as the LOD value and the proportion of phenotypic variance explained (PEV).

Trait		Condition	Year	Place	Identity	LOD	PEV	Marker	LG	Tresh[a]
RIP12	Flowering	glasshouse	2001	Cordoba	$QTL_{3A\text{-}RIP12}$	10.87	46.9	*FTa1*	3	3.1
	50% flower (initiation)	Field			$QTL_{3A\text{-}RIP12}$	4.52	22.0	*FTa1*	3	2.9
	Flowering (end)				$QTL_{end\text{-}RIP12}$	8.54	37.4	*COLh*	3	2.7
	Maduration				$QTL_{mat\text{-}RIP12}$	8.91	39.4	*COLh*	3	3.1
	50% flowering		2004		$QTL_{3A\text{-}RIP12}$	14.83	51.1	*FTa1*	3	2.9
					$QTL_{4B\text{-}RIP12}$	3.67	9.2	*STMS11*	4	3.3
			2008		$QTL_{3B\text{-}RIP12}$	6.34	29.6	*COLh*	3	2.9
			2014		$QTL_{3A\text{-}RIP12}$	8.48	29.8	*FTa1*	3	3.0
					$QTL_{4A\text{-}RIP12}$	5.36	17.3	*GAA47*	4	2.8
CRIL2	Days to first flower	LD, phytotron	2016	Hobart	$QTL_{3C\text{-}CRIL2}$	2.95	9.6	S1202p50545	3	2.8
					$QTL_{4C\text{-}CRIL2}$	2.56	8.3	S2032p76148	4	3.1
		SD, phytotron			$QTL_{3A\text{-}CRIL2}$	50.22	85.2	*FTa1*	3	2.6
RIP5	50% flower	Field	2003	Cordoba	$QTL_{3E\text{-}RIP58}$	9.68	38.7	*WRKY*	3	2.7
					$QTL_{3D\text{-}RIP5}$	3.04	8.7	*FTa1/2*	3	
				Mengibar	$QTL_{3D\text{-}RIP5}$	5.79	26.9	*FTa1/2*	3	2.8
RIP8	50% flower	Field	2003	Rep1	$QTL_{3E\text{-}RIP58}$	7.58	29.2	TA125	3	2.6
				Rep2	$QTL_{3E\text{-}RIP58}$	6.85	29.0	TA125	3	2.7

[a] Threshold LOD for a 0.995 confidence value, calculated through permutation test for each trait and linkage group.

CRIL2

Under LD conditions, only one minor QTL for flowering time ($QTL_{3C\text{-}CRIL2}$) was detected in the CRIL2 population, which explained 9.6% of the observed variance. This QTL was located at the top of chromosome 3, and most closely associated with marker S1202p50545 (Fig 4.6 A). There

was also evidence for the possible existence of a second QTL ($QTL_{4C\text{-}CRIL2}$) on chromosome 4 near marker S2032p76148explaining 8.3% of the variation (Fig 4.5, Table 4.5), but this did not quite reach the 0.995 confidence threshold for statistical significance.

When the same population was grown under 8h SD conditions, the results obtained were completely different. A single major QTL ($QTL_{3A\text{-}CRIL2}$) was identified on LG3, in a location distinct from that of the QTL under LD, between markers SUVH4 and *CDF2d* (Fig 4.6 A). The high proportion of phenotypic variation explained by this QTL (over 85%) means that it behaves essentially as a Mendelian locus and flowering time under these conditions is effectively under monogenic control. Within the $QTL_{3A\text{-}CRIL2}$ interval, the highest LOD value was obtained for the *FTa1* gene marker (Table 4.5), indicating a very strong linkage between this gene and flowering time in SD. In fact, a perfect association was found between the flowering response of the RILs and the *FTa1* marker genotype. The 55 RILs that were able to flower in SD all carried the ICC4958 allele at *FTa1*, and the remaining 69 non-flowering lines all carried the wild (PI489777) allele. A detailed analysis of the recombination in this region along with the competence of the RILs to flower narrowed the interval given by QTL analysis; no recombination was observed within the 850 Kb region delimited by markers SUVH4 and GATA9, which clearly indicates that this is the candidate region for flowering control in chickpea (Fig 4.4).

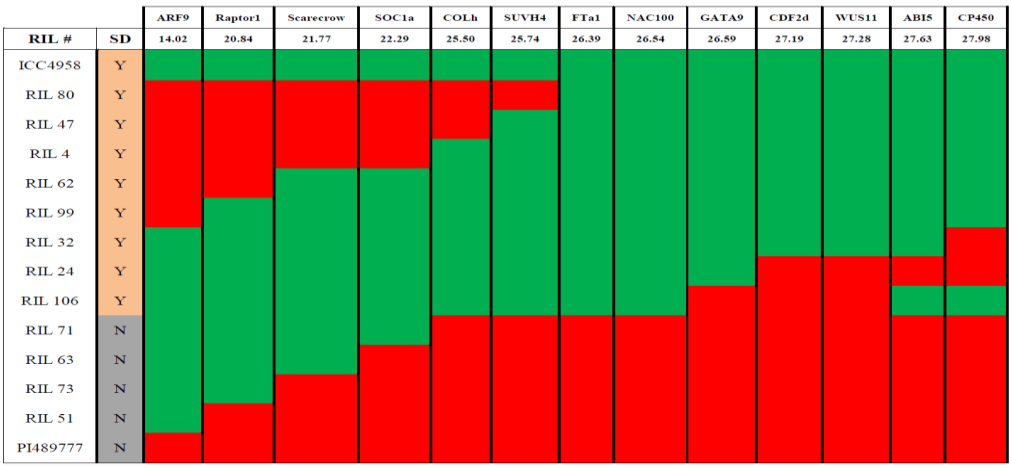


Figure 4.4 The recombination breakpoints, in a region of chromosome 3 spanning 13.96 Mb size, among 12 recombinant inbred lines from CRIL2. Parental genotypes are shown, and their colors indicates genotype of the RILs for the markers (ICC 4958 in green color, PI489777 in red color). Numbers under markers correspond to the physical position (in Mbp) of the markers in the chromosome 3, according to the genome of cultivar CDC Frontier available in NCBI. SD column indicates competence of the RILs to flower (Y) or not (N) under short photoperiod (8 h).

RIP12

QTLs for days to flowering in glasshouse and field conditions had been previously identified in LG3 by Cobos *et al.* (2009), based on data from 2001. To map these QTL more precisely with respect to the physical map, new markers within the interval were added and the analysis repeated. This analysis resulted in a new, narrowed interval defined by markers *SUVH4* and *CDF2d*, which is exactly the same as the interval identified for $QTL_{3A\text{-}CRIL2}$ in the CRIL2 population. This means that the same locus is likely to be acting in both populations, but since we have no definitive proof of this, the QTL found in RIP12 will be referred for further discussion also as $QTL_{3A\text{-}RIP12}$. Another similarity with the CRIL2 result is that the highest LOD scores (10.87 and 4.52 in glasshouse and field, respectively) were obtained for the *FTa1* marker (Fig 4.6 B). In 2001, $QTL_{3A\text{-}RIP12}$ accounted for 46.9% of the total phenotypic variance in flowering time in glasshouse and 22% in the field, and this large contribution was generally consistent in the 2004 and 2014 seasons (Table 4.5). In 2008, a QTL ($QTL_{3B\text{-}RIP12}$) for flowering time was also found in this general region ofLG3, but its position appeared slightly shifted relative to these others. This QTL was defined by the markers Q051828 and *FTa1*, with *COLh* as the peak marker accounting for 29.6% of the variance in the trait.

Days to end of flowering and days to maturity were also measured in the RIP12 field trial during the 2001 season. A major QTL ($QTL_{end\text{-}RIP12}$ and $QTL_{mat\text{-}RIP12}$) was found for each trait, with LOD scores of 8.54 and 8.91, respectively. In both cases, *COLh* was the peak marker, and the QTLs explained over 35% of the variation, showing that they had a big impact on the trait.

Chromosome 4 also seems to be involved in the control of flowering time in this population, as two secondary QTLs were identified in LG4 for days to 50% flowering in 2004 and 2014, over the markers GAA47 ($QTL_{4A\text{-}RIP12}$, LOD 5.36, PEV = 17.35) and STMS11 ($QTL_{4B\text{-}RIP12}$, LOD 3.67, PEV = 9.2%). The small distance between these two markers in RIP12 genetic map (3.6 cM) and the physical chromosome (0.82 Mbp) suggests that these two QTLs are probably due to the effect of the same locus (Fig 4.5).

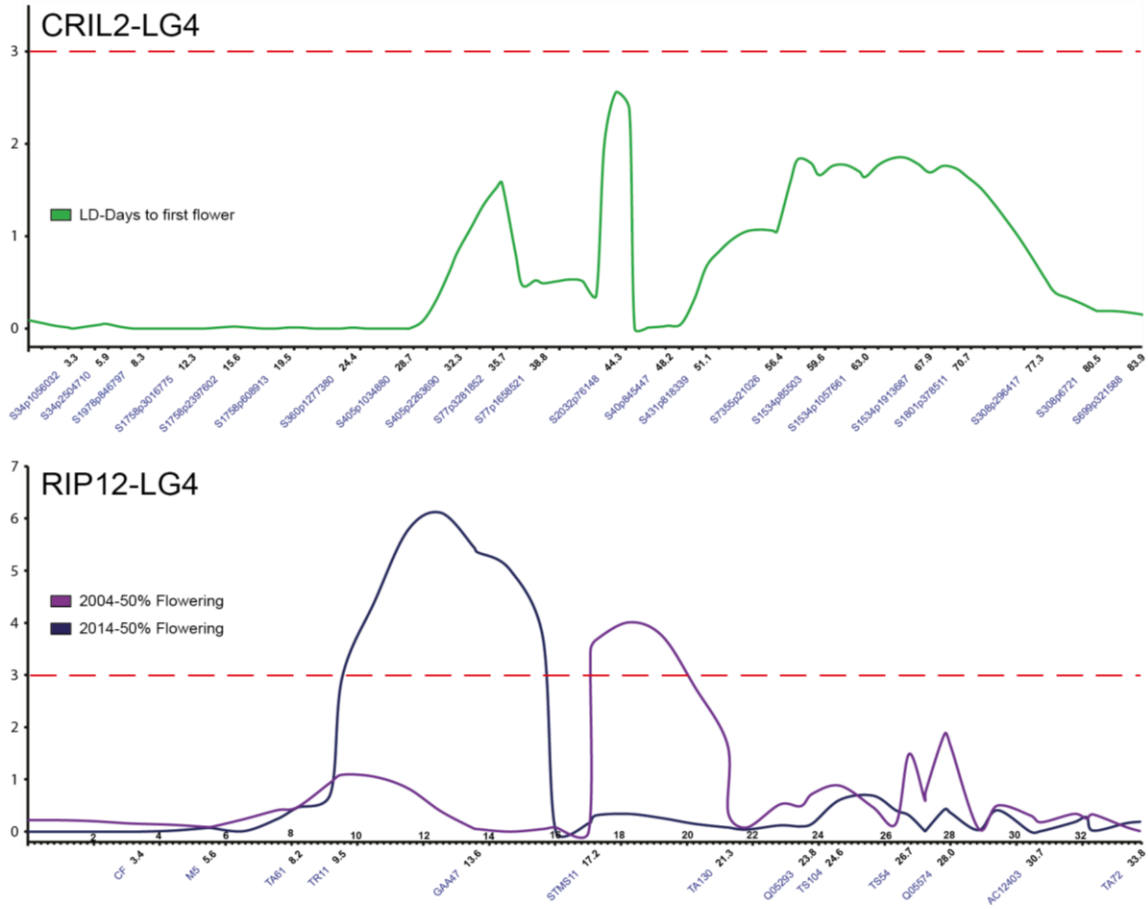


Figure 4.5 QTLs detected for flowering time in Linkage group 4 in the interspecific populations CRIL2 and RIP12.

RIP5/RIP8

In the intraspecific population RIP5 two different QTLs for days to 50% flowering were identified in the two locations analysed. Both were on chromosome 3, but only one of them ($QTL_{3D\text{-}RIP5}$), associated with the *FTa1* marker and delimited by markers TA142 and WUS11, was consistent across both environments. In Mengibar, $QTL_{3D\text{-}RIP5}$ was the only QTL detected, with a LOD of 5.79 and explained 26.9% of the variation in flowering time (Fig 4.6 C, Table 4.5), whereas in Cordoba this $QTL_{3D\text{-}RIP5}$ had a secondary role; it was less significant (LOD 3.04) and its effect on days to flowering accounted for only 8.7% of the observed variation. At the Cordoba site, a second QTL (QTL_{3E}, LOD 9.68) at a distinct position on chromosome 3 had a much stronger influence, explaining 38% of the observed variation (Table 4.5). This QTL was located between markers PRT6 and LOB189 and was most closely associated with markers WRKY and PRSP.

In the reciprocal cross population, RIP8, each replication of the trial was evaluated as an independent experiment, and gave very similar results (Fig 4.6 D). One major QTL in LG3 was found in each repetition, delimited by the same interval (H2I01-WRKY) and associated to marker TA125 in both cases. The proportion of phenotypic variance explained by TA125 (29.2 and 29%) and the LOD values (7.58 and 6.85) were also very similar between replications. This QTL co-locates with QTL_{3E}, and since RIP 5 and RIP 8 derive from reciprocal crosses, common QTLs between both populations are expected. Therefore, for the rest of this thesis they will be considered to be equivalent and referred to as $QTL_{3E\text{-}RIP58}$, delimited by markers PRT6 and LOB189. This interval is broader than the ones defining $QTL_{3E\text{-}RIP58}$ in each population separately, but it contains both. Using this conservative approach we ensure that the selected interval includes the gene responsible of $QTL_{3E\text{-}RIP58}$, which will be valuable for the future selection and targeting of candidate genes.

In summary, three regions have been associated with flowering time in chromosome 3 among two interspecific chickpea populations (Fig 4.7). One of them (corresponding to the interval of both $QTL_{3A\text{-}RIP12}$ and $QTL_{3A\text{-}CRIL2}$), linked to the *FT* cluster, was common to both populations and recurrent across environments. This region is likely to contain a major locus responsible of an important part of the phenotypic differences between both species. In the interspecific crosses evaluated, another two regions were found controlling flowering time. One of them ($QTL_{3D\text{-}RIP5}$) is also associated with the marker over the *FT* cluster, what suggest that the same locus acting between species can be a source of flowering variation within cultivated germplasm. However, its contribution to this trait is more variable and influenced by environmental condition.

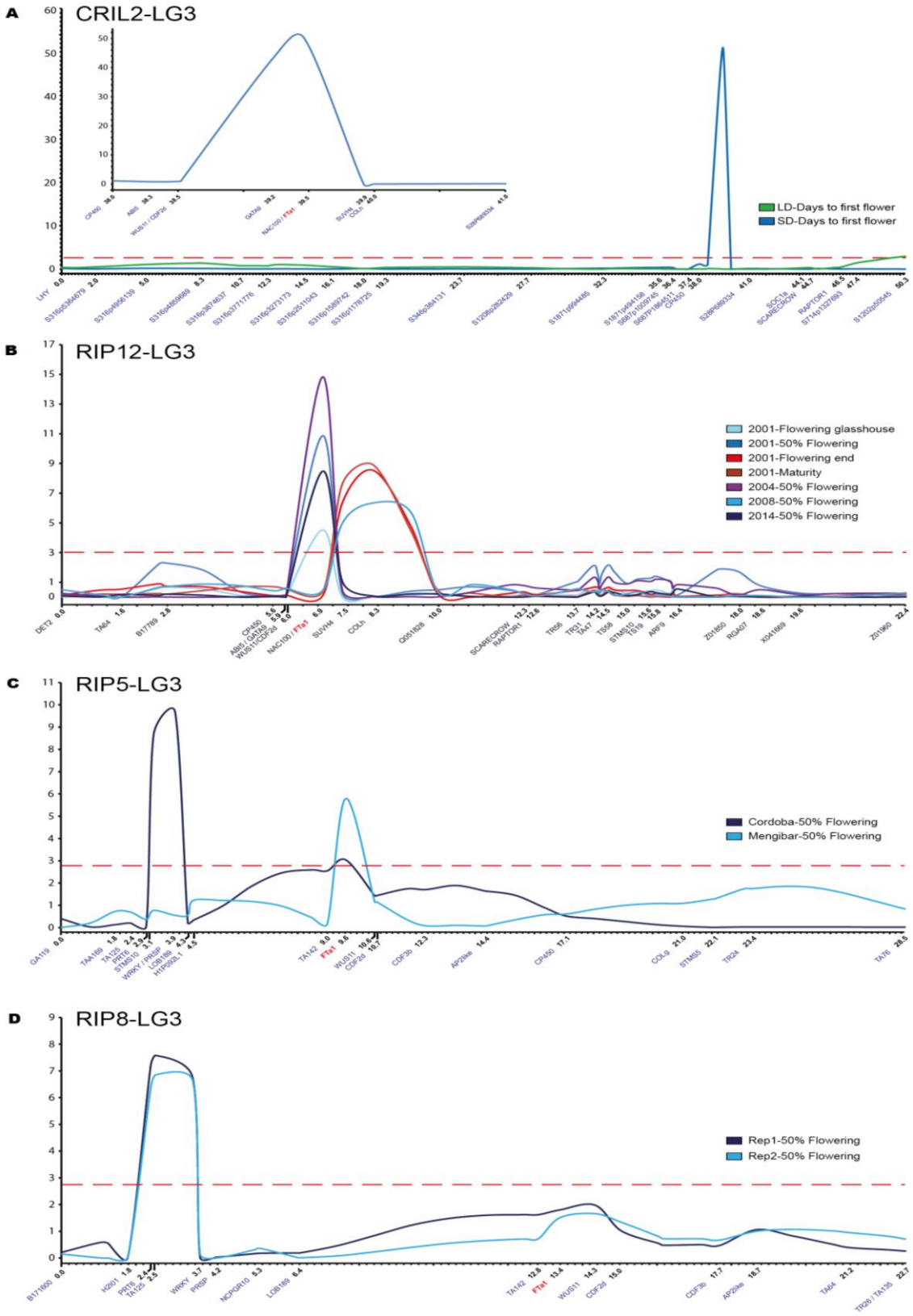


Figure 4.6 Flowering time QTLs detected in linkage group 3 in the four populations analysed.

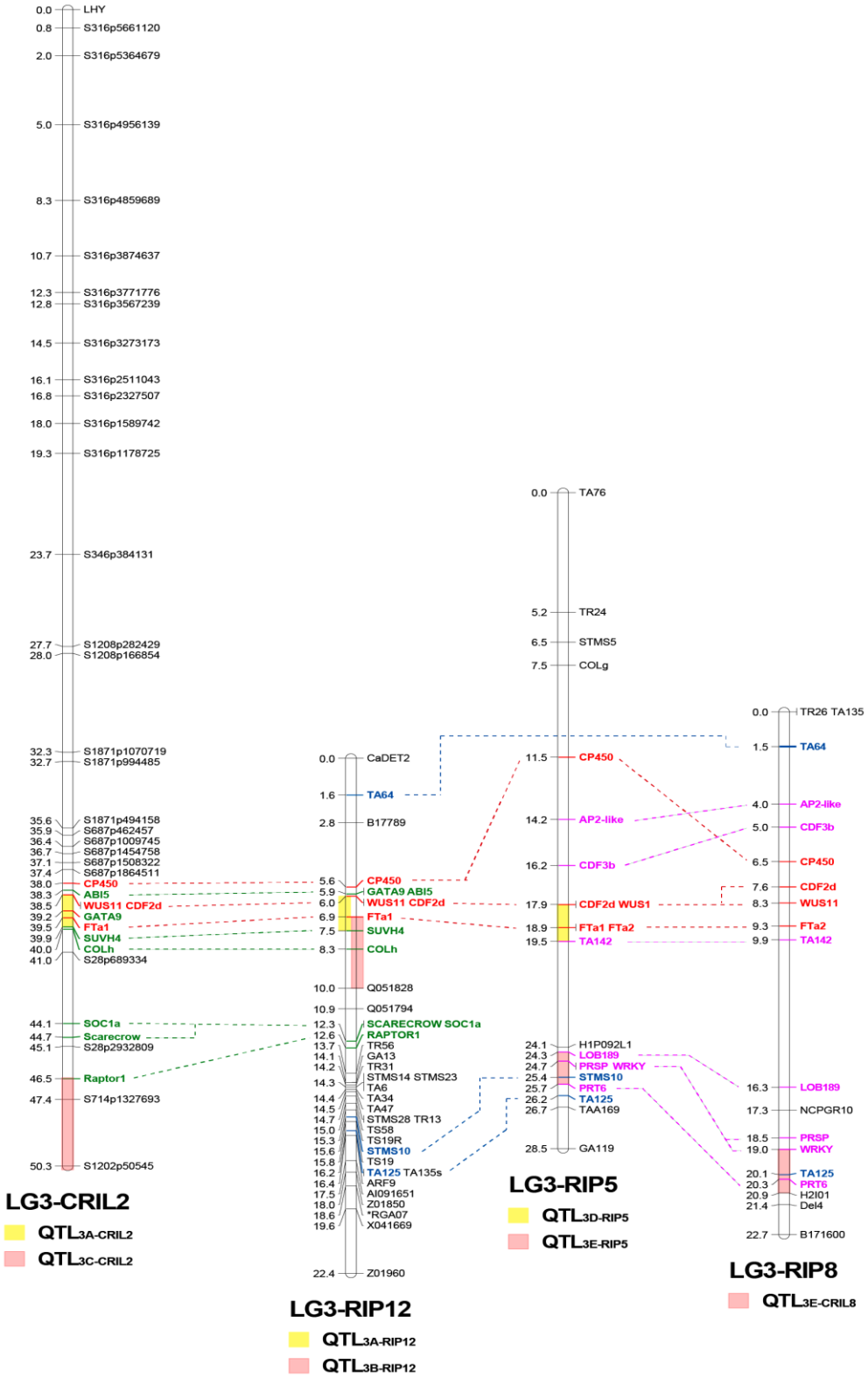


Figure 4.7 LG3 in the four chickpea populations, showing common markers and the co-location of the flowering QTLs obtained in this study (Table 4.5). Markers colour indicate that they are common to the 4 populations (red), to 3 populations (blue), to both intraspecific populations (pink) or to both interspecific populations (green).

4.3.4 Expression profile of chickpea FT genes

The results presented above have demonstrated the presence of QTL for flowering time centered on the *FTa1* gene in three different chickpea populations. In view of the important role known for *FTa1* in the promotion of flowering in pea and *Medicago truncatula* (Laurie et al. 2011; Hecht et al. 2011), one possible explanation for the action of this QTL is that it might interfere with regulation of *FTa1* and/or other members of the *FTa-FTc* cluster, causing elevated expression in plants carrying the domesticated (early) flowering allele. It was therefore of interest to compare expression of these genes between the parental lines of the populations in which the QTL was detected. The developmental expression of *FT* genes has not previously been examined in chickpea, and it was therefore also of interest to gain a general understanding of how *FT* genes and other flowering-related genes are regulated in chickpea.

In order to achieve these two aims, a time-series expression experiment was conducted in which the *FT* expression patterns from the six parental lines of the four populations were compared under both long and short photoperiods. As previously stated, CRIL2 is a reference chickpea population. Therefore, the time-series experiment in the parentals of this population (ICC4958 and PI489777) was designed to be longer than the rest (12 weeks vs 4 weeks), and the expression of additional flowering-related genes was measured, in order to gain deeper insight of the floral regulation pathways in chickpea.

Under the controlled conditions used in this study, all the early parents of the four populations had a similar flowering time, of around 5 weeks after emergence under LD and 7 to 8 weeks under SD. In comparison, the late parents took 8-10 weeks to flower in LD and did not flower for more than 12 weeks under SD.

Orthologs of the *A. thaliana* inflorescence identity gene *APETALA1* are known as *PROLIFERATING INFLORESCENCE MERISTEM* (*PIM*) genes in pea and Medicago, and their initial induction in the shoot apex has been used as an indicator of flower commitment (Laurie et al. 2011; Hecht et al. 2011). The expression of the chickpea *PIM* ortholog was induced in apical buds at least one week prior to the visible appearance of flower buds in all six parental genotypes (Fig 4.8), confirming that this gene is also an indicator of the vegetative to reproductive switch in chickpea. The apical expression of another floral identity gene *UNIFOLIATA* (*UNI*), the ortholog

of Arabidopsis *LEAFY* (*LFY)*, followed the same pattern when examined in CRIL2 parent lines (Fig 4.10). Therefore, only *PIM* will be used routinely an indicator of flowering in the other populations.

Expression of *FTa1* was detected in both leaf and apex tissues, although its expression was one order of magnitude higher in leaves (Fig 4.8, Fig 4.9). While no significant differences were found for *FTa1* levels in the shoot apex between early/late accessions or photoperiod treatments, its pattern of regulation pattern in the leaf can be correlated with the induction of *PIM* in the apex and it can thus can be considered as a good candidate to encode a florigen capable of moving from leaf to apex to induce flower development.

FTa2 was also expressed in both tissues and, as for *FTa1*, expression in leaves was higher than in in apical buds. Interestingly, wild and domesticated accessions showed differential expression of this gene. In the wild species *C. reticulatum* (accessions PI489777 and Cr5-9), *FTa2* induction occurred gradually, especially in apex tissues, and can be correlated with flowering (Fig 4.8). In comparison, the expression in cultivated *C. arietinum* (ICC4958 and ICCL81001) was highly upregulated in both tissues, even in one week old seedlings. In both genotypes, *FTa2* transcript level was similar in LD and SD (Fig 4.8, Fig 4.9) indicating that the differential response of cultivated *FTa2* to photoperiod may. However, there are evident differences among accessions; ICC4958 shows the highest *FTa2* level, followed by ICCL81001, ILC3279 and finally both *C. reticulatum* accessions. Surprisingly, no expression of *FTa2* was detected in WR315, the early parent of RIP5/8, and this was subsequently discovered to reflect a deletion of this genomic region (See Chapter 6).

Transcription of *FTb* was negligible in apex samples. In leaves, it was up-regulated under LD in all six chickpea lines tested, but was not detected under SD, and thus shows a clear and strong regulation by photoperiod. Chickpea *FTc* was expressed only in the apex during the phase where flower commitment takes place. Its expression mirrored that of the meristem identity genes *PIM* and *UNI*, so it is a potential regulator of floral transition. The longer experimental time course for the CRIL2 parents revealed that it is also expressed in leaves at a later stage of plant development after flowering has already taken place (Fig 4.10), suggesting that this gene could be involved in a post-flowering process such as pod formation.

Recently, a third member of the *FTa* subclade, *FTa3*, has been described in temperate legumes (see chapter 3, Section 3.3.1) but nothing is yet known about its regulation or role. The developmental expression profile of chickpea *FTa3* was only examined in the CRIL2 parents, to test its involvement in flower development. Expression was undetectable in apex, and in leaf, only occurred late in plant development once flowering had already happened, resembling the pattern obtained for foliar *FTc*.

Apical expression of *TFL1b* and *TFL1c*, chickpea homologs of the *A. thaliana* floral repressor *TFL1*, was also measured in the CRIL2 parents. The expression of both genes followed a similar pattern, although only in the case of *TFL1c* the expression levels were high enough to be compared with confidence. Expression is elevated in *C. reticulatum* under non-inductive conditions (SD), while under LD and in the cultivated accession the expression level of both genes gradually drops with aging (Fig 4.10). This profile was opposite to that obtained for the *FT* genes and thus fits to that of a flowering repressor.

The important soybean regulator *E1* is an apparently legume-specific gene that can also influence flowering regulation in Medicago, according to Zhang et al. (2016). One homolog was found in chickpea. Its expression shows a spatial pattern restricted to leaf. No differences between accessions or photoperiod treatments were found, so its role in flowering induction in chickpea seems unlikely in the conditions studied.

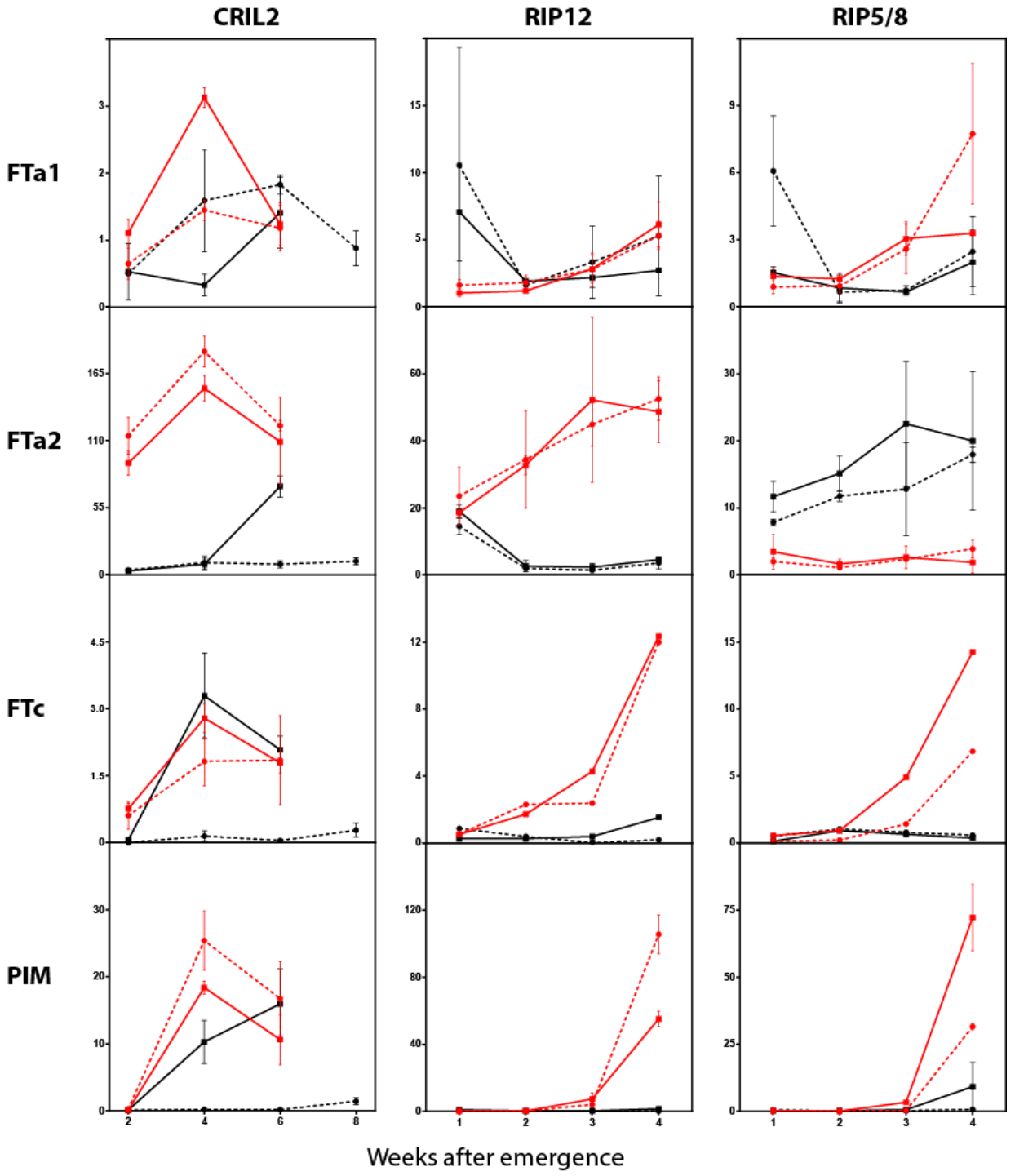


Figure 4.8 Expression profiles in apex of *FT* genes and the floral indicator *PIM* in the parental lines of three chickpea populations. Early flowering parents are represented in red (ICC4958, ICCL81001 and WR315 for CRIL2, RIP12 and RIPs 8/5, respectively) while late parents are shown in black (PI489777, Cr5-9 and ILC3279). Continues lines and squares indicate long day and dashed lines/circles short day photoperiod. The average ± SE of 2 biological replicates is shown, and transcripts were normalized against *ACTIN*.

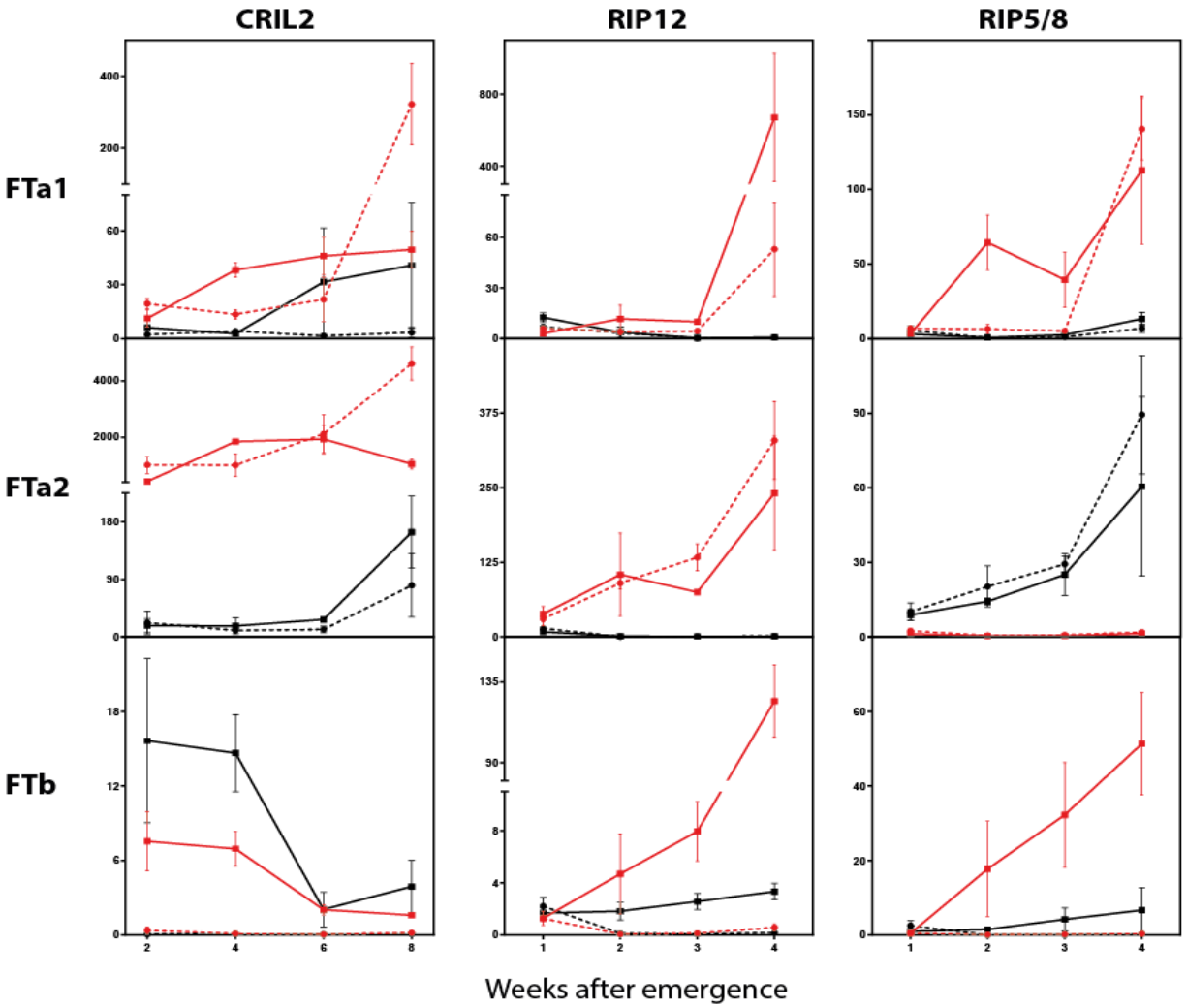


Figure 4.9 Expression profiles of *FT* genes in leaves in the parental lines of three chickpea populations. Early flowering parents are represented in red (ICC4958, ICCL81001 and WR315 for CRIL2, RIP12 and RIPs 8/5, respectively) while late parents are shown in black (PI489777, Cr5-9 and ILC3279). Plain lines and squares indicate long day and dashed lines/circles short day photoperiod. The average ± SE of 2 biological replicates is shown, and transcripts were normalized against *ACTIN*.

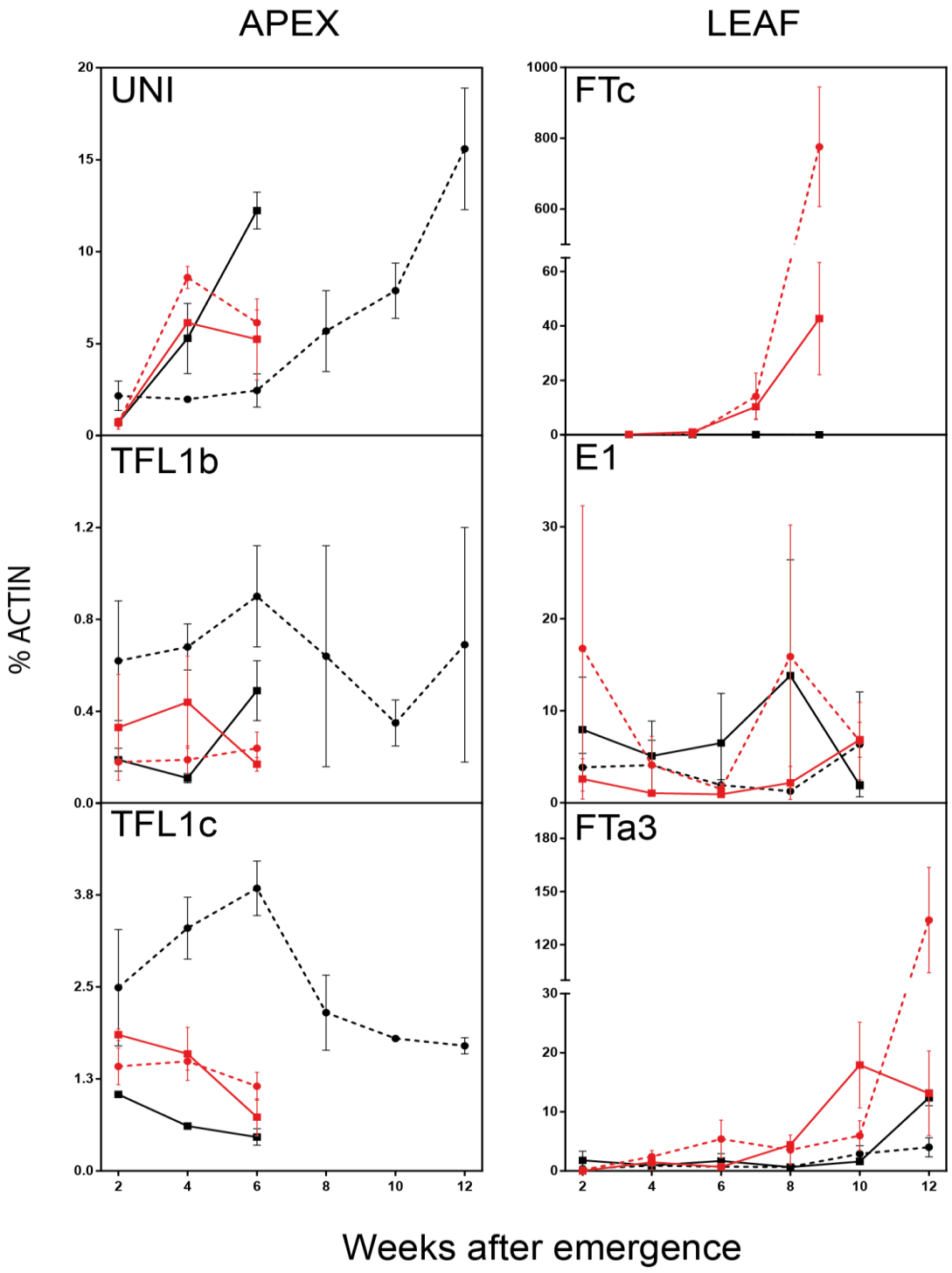


Figure 4.10 Relative expression of flowering-related chickpea genes in CRIL2 parental lines. ICC4958 and PI48977 are represented by red and black lines, respectively. Continuous lines/squares indicate long day condition while dashed lines/circles represent short day. *UNI*, *TFL1b* and *TFL1c* expression was measured only in apex. E1 and *FTa3* expression was measured in both tissues but found only in leaf. β-actin gene used as reference. Primers and conditions can be found in Table 4.3.

4.4 Discussion

Flowering time is a complex trait and work in model species suggests that it involves the interaction of hundreds of genes in an intricate network, and variation in most of these genes can lead to an altered phenology. The large number of flowering-related QTLs reported in chickpea and the involvement of all eight chromosomes is therefore not surprising particularly considering the potential contribution of major and minor genes acting in different crosses and responding to specific environments (reviewed in Table 1.2). Of special relevance are the major QTL with high PVE values, and those overlapping in certain location of the chromosomes (Fig 3.5), indicating the presence of a locus with an important potential to influence flowering time. Across all reported studies that have examined flowering-related phenotypes, chromosomes three and four are most prominent with more than 40 QTLs. Identifying the underlying genes would therefore help to understand and manage a significant component of flowering time variation. This is particularly important in the case of genes that confer early flowering, as earliness is a desirable trait used by breeders as an escape mechanism from abiotic stresses that typically constrain chickpea yield in traditional environments (Kumar and Abbo 2001; Subbarao et al. 1995; Anbessa et al. 2006). The significant continuing effort to find and map sources of earliness in chickpea reflects the significance of this trait in chickpea improvement (Raghu 2001; Bonfil et al. 2006a; Anbessa 2006; Mallikarjuna et al. 2017).

In particular, the specific interest in a cluster of QTLs in the central region of chromosome 3 has been outlined in chapters 1 and 3. A detailed analysis of published marker data in relation to the chickpea physical map identified several candidate genes in this region, among which a cluster of three *FT* homologs (*FTa1*, *FTa2* and *FTc*) have been proposed as the most plausible (Chapter 3; Weller and Ortega, 2015). The main objective of this chapter was to examine this hypothesis - firstly by generating a detailed, gene-based map of the of the QTL region to refine the position of the *FT* cluster relative to the QTL, and second, by examining whether expression of these candidates might differ between the early and late parents of each population.

QTLs in LG3 in interspecific crosses

Photoperiod and temperature are the most important environmental factors that influence flowering time in plants and both factors appear to be important in control of chickpea flowering

(Daba et al. 2016b; Roberts et al. 1985; Berger et al. 2011). A significant influence of photoperiod is evident when comparing the flowering response of all parental lines. All lines flower earlier under LD than SD, and differences between lines are much smaller. For example, flowering in ICC4958 and PI489777 was quite similar under LD, but delayed to a quite different extent under SD. The lack of phenotypic difference under LD can explain why only one minor QTL was identified under these conditions. The fact that all plants flowered early and in a short window of time under LD shows that the photoperiod experienced by the plants in this condition was sufficient to override any other genetic difference, and it can be assumed that all lines have an intact photoperiod-response pathway. The means in flowering date obtained for both parental lines of CRIL2 under LD are very close to that obtained by Daba et al. (2015) in a collection of chickpea accessions grown using similar conditions, indicating that these conditions are enough to fully activate the LD response.

The difference in flowering responses of the CRIL2 parents was particularly striking under an 8h SD photoperiod. Cultivated chickpea flowered at 60 days whereas wild chickpea was not able to flower for more than 130 days, and seemed unlikely to flower at all. QTL analysis showed that this difference observed in the flowering behaviour of ICC4958 compared to PI489777 is mostly attributable to a single gene in the central region of LG3 delimited by markers SUVH4 and CDF2d (1.4Mbp). Fine mapping delimited the $QTL_{3A\text{-}CRIL2}$ to a 850 kbp region between markers SUVH4 and GATA9 containing 59 genes (appendix 4.6), among which the *FTa1*/*a2*/*c* genes were identified as the most closely associated candidates. In fact, a marker for the *FTa1* gene showed perfect correlation between the domesticated allele and the ability to flower (Fig 4.4). A major QTL defined by exactly the same interval (SUVH4-CDF2d) was also found in another interspecific population (RIP12), consistent across almost all environments. *FTa1* was also the marker explaining most of the observed differences in flowering date within this population (up to 51%).

The fact that this major QTL was found in two different interspecific populations where the domesticated parents were unrelated suggest that it may have occurred relatively early in the history of chickpea cultivation and may even have been involved in its domestication. This is difficult to conclude definitively from analysis of only two populations. However, two other interspecific crosses are described in the literature, and both also report a major QTL for

flowering time on chromosome 3. In a cross between the *desi* accession ICC3996 and the wild ILWC1847, Aryamanesh et al. (2010) found two QTL on chromosome 3 that together explained most of the observed variation (90%). One of these was located in the exact interval reported initially in RIP12 by Cobos et al. (2009), flanked by markers TA64 and TA142. A second, more recent study involving an interspecific cross also found a major flowering time QTL on chromosome 3 that was consistent across different years and environments (Das et al. 2015b). However this study also used ICC4958 as the domesticated parent, and therefore, although consistent with results obtained in CRIL2, it does not add independent support to the argument.

QTLs in LG3 the intraspecific crosses

QTL analysis carried in the intraspecific populations identified two major regions controlling flowering time in chromosome 3. One of these, termed $QTL_{3E\text{-}RIP58,}$ was located within an interval of 4 Mbp defined by markers PRT6 and LOB189 and explained a substantial proportion of the observed variation for flowering time (29-38.7% PEV). A comparison with other published studies suggests the presence of a QTL in a similar position in other intraspecific crosses. For example Hossain et al. (2010), in the cross between the *desi* cultivar ICC3996 and the *kabuli* line S95362, reported a QTL between markers TR56 and TS19, a genomic region that includes the $QTL_{3E\text{-}RIP58}$ region. Daba et al. (2016a) also reported a minor QTL associated to the marker CAV1SC48.1P396061, which is closely enough (1.1 Mbp) in the chromosome to the $QTL_{3E\text{-}RIP58}$ interval to at least consider the possibility that they might represent the same locus. Finally, Hamwieh et al. (2013b) also found a QTL consistent across different seasons and environments on chromosome 3. However, the associated markers H6C07 and H4G07, are located on the upper part of the chromosome, considerably above the flanking markers for the other QTL, making it unlikely this QTL is the same as the one identified in RIP5/8 and in the other studies. The question of the molecular identity of the gene under QTL_{3E} needs to be addressed. Unlike the adjacent QTL_{3A}, for which several candidate genes were identified, none of the flowering-related genes identified in Chapter 3 were located in the $QTL_{3E\text{-}RIP58}$ interval. A total of 244 genes are annotated in this interval according to NCBI (appendix 4.7), and some of them are worth to mention, as they have biological functions that could potentially be connected to flowering. These include an F-BOX protein (GeneID 101493711), which belongs to a big protein family that has been shown to participate in the control of several key biological

processes in plants, including flowering (Kipreos and Pagano 2000; Durfee et al. 2003; Kim et al. 2013a; Wang et al. 2002). Several hormone-related genes are also present in this interval, including genes encoding an abscisic acid receptor, auxin transporter and the *Ethylene-overproduction protein 1*, which might have the potential to influence the plant life cycle and so potentially lead to altered phenology (Wasilewska et al. 2008; Wang et al. 2013; Shu et al. 2016; Chae et al. 2003; Ogawara et al. 2003). Finally, the extense WRKY family of transcription factors (Wu et al. 2005; Zhang and Wang 2005) also have potential to explain the phenotypic differences as proven in other species; in rice, the late flowering phenotype of the *dlf1* mutant is caused by an insertion in *OsWRKY11* (Cai et al. 2014). In Arabidopsis, mutant analysis demonstrated that *WRKY71* accelerates flowering by inducing the expression of floral promoters such as *FT* and *LFY* (Yu et al. 2016). Two more members of this family have been reported to have opposites effects in flowering time of Arabidopsis under short photoperiod; *WRKY12* promotes flowering whereas *WRKY13* delays it, in a mechanism that seems to involve the hormone gibberellin, according to (Li et al. 2016). Two WRKY proteins can be found in the $QTL_{3E\text{-}RIP58}$ interval and interestingly, one of them was targeted in the present study (GeneID 101511519) and was the marker with a higher LOD value in $QTL_{3E\text{-}RIP58}$ when detected in RIP5. Finally, it may be that the causal gene is novel, with a previously unknown role in flowering time control. In any case, it is likely that future work including fine mapping will be needed to identify the responsible gene.

A second intraspecific $QTL_{3D\text{-}RIP5}$ was detected only in RIP5, despite the fact that both intraspecific populations derive from reciprocal crosses. Although not significant in RIP8, the LOD curve shows a rise between markers TA142 and CDF2d (the region of the $QTL_{3D\text{-}RIP5}$), indicating a quasi-significant association of this region with flowering (Fig 4.6 D). In any case, the region identified by $QTL_{3D\text{-}RIP5}$ seems to be involved in the control of flowering in other crosses as reflected by previous QTL reports. Firstly, it overlaps with the major QTL found in the interspecific crosses, described above. However, a QTL in this region has also been reported from crosses within domesticated chickpea. For example, Mallikarjuna et al. (2017) found 4 QTLs from 3 different crosses overlapping in the region between TA142 and TA64, which corresponds to a large portion of chromosome 3 and includes the *FT* cluster. Rehman et al. (2011) reported another QTL in crosses with ILC3279 as common parent that was flanked by markers TA6 and NCPGR12. This region includes QTL_{3D}, so they can be pointing to the same locus.

QTLs in LG4

Numerous studies have reported significant flowering-related QTLs in LG4 (Cobos et al. 2007; Rehman et al. 2011; Varshney et al. 2014b; Pushpavalli et al. 2015; Upadhyaya et al. 2015; Samineni et al. 2016; Gowda et al. 2011; Vadez et al. 2012; Daba et al. 2016a), and it has become well known as the site of a 'QTL hot-spot'' (Varshney et al. 2014b). In the present study, two QTLs on this chromosome were identified in the interspecific population RIP12, associated most closely with markers STMS11 and GAA47. Whether these two QTLs are independent or they represent the same locus is uncertain, although the adjacent position of these two markers in the linkage map makes the second scenario more likely. The low marker density in the neighbouring region can lead to imprecision in positioning the QTL interval across environments, so future work involving more markers in this region is needed to stabilise the QTL and obtain a more robust estimate of its position. In any case, as discussed in chapter 3, both GAA47 and STMS11 are included in a larger interval of chromosome 4 defined by markers NCPGR80 and NCPGR127 (3.4 to 16.8 Mbp in the physical map) that seems to host at least one gene playing a major role in the flowering network, according to association studies in both inter- and intraspecific crosses (Mallikarjuna et al. 2017; Cobos et al. 2007; Pushpavalli et al. 2015; Daba et al. 2016a; Varshney et al. 2014b). Chickpea homologs of some well-known Arabidopsis flowering time regulators such as *GIGANTEA* (*GI*), *TEMPRANILLO* and *ELF6* can be found in this region, as discussed in Chapter 3, and at least in the case of GIGANTEA, there is already evidence that it has an important role in legumes (Ridge et al. 2016; Watanabe et al. 2011; Hecht et al. 2007b; Liew et al. 2009).

Maturity

Traditionally, breeders have used days to flowering as an indicator of earliness in the duration of the crop (Kumar and Abbo 2001). A correlation of days to flowering and days to maturity has been reported in several studies across different crosses and environments (Anbessa et al. 2007; Das et al. 2015b; Hamwieh et al. 2013b; Monpara and Dhameliya 2013; Gaur et al. 2014a) indicating either a linkage between the two traits and/or the existence of genes with an effect on both traits. The existence of cultivars with early flowering and late maturing and *vice versa* suggests that these traits can be separated to some extent (Summerfield and Roberts 1988). In the present study, QTLs for days to flowering and days to maturity in RIP12 were very close to each

other but clearly in different positions (Figure 4.6 B), suggesting that in this population these two traits may be regulated by different genes. This evidence provides the first indication that flowering time and maturity in chickpea might be separable at the genetic and molecular level.

The marker *COLh* was the peak marker for both end of flowering and days to maturity QTL, and its physical position is very close to the $QTL_{3A\text{-}RIP12}$ governing flowering (Fig 4.6 B), with their peak markers separated by only 0.8 Mbp. Rehman et al. (2011) found a QTL in LG3 involved in the control of flowering and maturity between markers TA6 and NCPGR12, which includes the interval of $QTL_{mat\text{-}RIP12}$ obtained in this study, supporting the association of this region of chromosome 3 with the control of both traits. *CONSTANS* is a key gene promoting flowering in Arabidopsis, but this role does not seem to be conserved in other Arabidopsis *CO*-like genes or in any of the *CO* genes in pea or Medicago (Wong et al. 2014). Nevertheless, the legume *CO* family, and specifically *COLh*, has not been exhaustively characterized, and this result does leave open the possibility that *COLh* could have a role in regulating maturity independent of flowering induction.

Expression profile of the FT family

Previous reports have shown that the *FT* family in legumes is more complex than in Arabidopsis, comprising 5 or 6 genes depending that can be classified in 3 groups (*FTa*, *FTb* and *FTc*) (Hecht et al. 2011; Laurie et al. 2011). In Chapter 3, we characterized the PEBP gene family in chickpea, finding that the *FT* clade is composed by 5 members, 3 belonging to *FTa* (*FTa1*, *FTa2* and *FTa3*) and one to each of the *FTc* and *FTb* subclades. Here, we characterized the regulation patterns of all chickpea *FT* genes, and show that they have different expression profiles in response to environmental cues.

A primary objective of the expression analysis was the evaluation of the three *FT* genes in the cluster to determine the most suitable candidate. By definition, *FT* genes promote flowering. Thus, if any of these genes were causing the early phenotype of *C. arietinum* compared to *C. reticulatum*, we would expect to see a gain-of-function of that gene rather than a loss of function, since then a late phenotype would be expected, consequent with that of *ft* mutants in other species (Matsoukas 2015). According to legume studies on the subject, *FTa1* would be the most likely candidate, as it plays a key role in regulation of flowering in both pea and Medicago

(Hecht et al. 2011; Laurie et al. 2011), where its expression is upregulated by both photoperiod and vernalization. Results obtained in the present study are similar to those in the mentioned works, with an *FTa1* regulation pattern resembling the expected for a classic florigen, induced in leaf and moving to apex to stimulate the expression of the floral identity genes that control floral development. In pea, *FTa1* seems to promote flowering also under SD (Hecht et al. 2011), assuming the role of Arabidopsis *TSF* (D'Aloia et al. 2011; Yamaguchi et al. 2005).

FTc can also participate in floral induction, as its regulation in apex correlates well with flowering conditions, consistent with those described in pea and Medicago. Although its expression level is low, complementation studies suggest that is the most effective FT protein in the induction of flowering, suggesting that even a relatively low level of transcript might be sufficient for its function.

Very different *FTa2* expression patterns were found at an interspecific level; while the profile in *C. reticulatum* is quite low and consistent with that reported in related species, 3 out of 4 cultivated chickpea accessions showed a high level of transcription, with a pattern resembling a constitutive gene highly expressed at all times points. However, the complete lack of expression for this gene in accession WR315 clearly indicates that expression of *FTa2* is not required for flowering induction.

Resuming, we found that the timing and differential expression levels of both *FTa1* and *FTc* fit with their role as candidates, but unfortunately we can't draw any further conclusion regarding this issue.

Due to its central role in the flowering network, *FT* is subject to tight regulation (Andrés and Coupland 2012; Song et al. 2015; Amasino and Michaels 2010). Many factors (both promoters and repressors of its expression) bind to cis-elements in the regulatory regions of *FT*, so a mutation in any of these places blocking the action of a repressor transcription factor could cause a gain-of-function of these genes and an earlier expression like the one obtained here. However, as above indicated, many factors regulate FT expression, and there is still the possibility that the differences observed in *FTa1*/*FTc* expression between early and late parentals are due to the action of a trans-element. This is unlikely to be happening, as no other major QTL was found in any of the populations, but this alternative scenario cannot be discarded.

FT proteins and their homologs plays a conserved, universal role in the control of floral transition (Matsoukas et al. 2012). This family have been characterized in the legume species Medicago and pea, and information about their transcriptional regulation and interaction under different photoperiods and age is available (Laurie et al. 2011; Hecht et al. 2011). Therefore, as a secondary objective of the expression analysis in the present study, we tested the conservation of this model in chickpea. Overall, our results were consistent with these reports, including the above described regulation patterns of the *FT* genes in the cluster and a possible cross-regulation of different *FT* genes (Hecht et al. 2011). The expression of chickpea *TFL1* homologs, in particular that of TFL1c, with opposite pattern to *FT* and higher expression in the late parent under non-inductive condition is also similar to that observed in previous studies and fits a possible role as flower repressor.

In chapter 3, we reported that only one *FTb* gene was found in chickpea, while other legumes in the same clade present two *FTb* paralogs, namely *FTb1* and *FTb2* (e.g. Medicago and pea). The high level of homology among *FTb* proteins makes it difficult to discern whether the remaining one is an *FTb1* or *FTb2* ortholog (appendix 3.25). The photoperiod-regulated expression pattern obtained in this study is consistent with the behaviour of *FTb2* reported in other legumes (Hecht et al. 2011; Laurie et al. 2011), so we can assume that the only *FTb* present in chickpea is an *FTb2* ortholog.

A novel *FT* homolog, *FTa3*, was also identified in chapter 3. Since it has not been characterized in any other species, no information is available about its role or regulation. Its expression was restricted to leaf, with a maximum only occurring weeks after the commencement of flowering, so it seems to be irrelevant for flowering commitment under the conditions used in this study. Further work on this homolog will clearly be necessary to understand its function and regulation.

4.5 Conclusion

This chapter investigated the role of the *FT* cluster as candidate genes using molecular mapping and gene expression analysis. Co-location of $QTL_{3E\text{-}RIP5}$ and QTL_{3A} evidence that they correspond to the same gene, and in both cases the marker targeting the *FT* cluster was the most significantly associated locus. These observations suggest the presence of a locus in this region able to influence phenology at both inter- and intra-specific level. Although only a small number

of studies have used interspecific populations, they all agree on the relevance of this region for flowering induction, so here we propose that it should be considered as harbouring a potential domestication locus. Expression analysis shows an early induction of *FTa1* (in leaf) and *FTc* (in apex) in the early-flowering parent of both interspecific and intraspecific crosses, not only in LD but also under SD. These results support these *FT* genes as strong candidates for flowering time QTL in this region, although they should be taken cautiously in view of the fact that as *FT* genes are likely to be targets of different pathways, their expression could in principle also be influenced by other genomic regions. In particular, another major locus was found in both intraspecific populations, suggesting that the role of the chromosome 3 region may be shared with other loci in *C. arietinum*.

Chapter 5. Genetic control of branching and growth habit

5.1 Introduction

Crops are very different to the wild species from which they derive, to the point that in extreme cases their relationship is difficult to recognise. The accumulated set of changes imposed on a species by human selection is referred to the domestication syndrome, and includes major changes in features that facilitate human handling and harvest (such as indehiscent pods), phenology, organoleptic properties, and also, plant architecture, which has changed dramatically in many species (Huyghe 1998; Ward and Leyser 2004).

Branching pattern (understood as the number, location and length of branches), growth habit (defined as the angle of the branches in relation to the vertical axis) and determinacy are the main determinants of plant architecture. Collectively these traits have a great influence in many aspects of a plant life that are important for local adaptation, including light interception, water loss and susceptibility to disease and damage, and thus have been a common target in the domestication of most crops (Doebley et al. 2006). Cultivated chickpea has generally an erect or semi-erect phenotype, and resembles a small bush with primary branches held upright at an acute angle to the main stem. By contrast, wild species of chickpea have a prostrate growth habit with branches growing horizontally. This phenotype is genetically dominant over the erect habit and is also associated with a lack of apical dominance and an elevated number of branches (Ali et al. 2015; Singh and Shyam 1959; Aryamanesh et al. 2010). The erect growth habit of *C. arietinum* was selected during its domestication as a desirable trait that facilitates harvesting, a factor that has become even more important in recent times, especially in developed countries where this task is mechanized. As discussed in Chapter 1, cultivated chickpea has low genetic variability, and introgression of beneficial genes from wild *Cicer* species is likely to be an important approach in future improvement through classical breeding. However, this approach also means that undesirable traits from wild must be avoided, and knowledge about the genetic regulation of these traits can help in this process. Typical examples of beneficial genes from wild *Cicer* species include those specifying resistance to pests and abiotic stresses, but it has also been shown that genes for desirable yield components such as high number of fruiting branches and pods per plant can also be derived from wild relatives (Singh et al. 2005; Singh and Ocampo

1997). However, as mentioned above, these desirable branching components are associated with prostrate growth habit of *Cicer reticulatum*, which has generally been avoided in breeding programs, even though it can have a positive effect on yield, as proven in Mediterranean environments (Siddique and Sedgley 1985; Rubio et al. 2004). If the genetic basis for these two traits can be better understood and separated, it may assist in the development of new cultivars with a novel architecture that could potentially improve yield, and improve ability to select for desired architecture phenotypes in wide crosses.

5.1.1. Growth habit

In chickpea, the term growth habit is used by researchers essentially to indicate the angle that the branches form with the vertical axis or the ground. In general, the control of branch angle is not well understood but it appears to be a complex trait governed by multiple genes (Benlloch et al. 2015), and little is known about its molecular regulation. Two loci, *Hg* and *Hg2*, controlling growth habit in chickpea have been described to date. *Hg* was found to be linked to the enzyme 6-PGD on chromosome 3, whereas *Hg2* was mapped to chromosome 1 (Ali et al. 2015; Muehlbauer and Singh 1987; Kazan et al. 1993; Winter et al. 2000). A high degree of variation in growth habit has been found among cultivated accessions (Upadhyaya et al. 2006), so it is likely that a larger number of genes contribute to control of the intraspecific variation. Supporting this, a recent genome-wide association study (GWAS) using a collection of domesticated chickpea lines corroborated the importance of the above mentioned linkages groups (LG) in the control of growth habit but also described additional new loci on chromosomes 4 and 7 (Upadhyaya et al. 2017).

5.1.2 Branching

While the angle at which branches are held is a critical component of plant architecture, the number, length and origin of branches are also important factors. As in the case of flowering, branching is a very complex trait regulated by different pathways involving endogenous and environmental cues that converge to common integrators that control the decision of whether an axillary bud remains dormant or grows out to produce a branch (Rameau et al. 2015; Leduc et al. 2014; McSteen and Leyser 2005). Branch development is a well-characterized process at the morphological level, and in view of its importance for yield, its physiological and molecular

regulation is well studied in model species (Rameau et al. 2015) and diverse crops such as cereals (Hussien et al. 2014; Liang et al. 2014) and legumes (Yang et al. 2017; Nelson 1996; Beveridge et al. 2003; Sorefan et al. 2003). In chickpea, however, the control of branching has not received much attention and is not well understood. However, recent advances in chickpea genomic tools have provided new approaches to the study of its molecular regulation. In the past decade, several QTL and genome-wide association (GWA) analyses agreed and resulted in the identification of several genomic regions and candidate genes controlling the number of branches (Varshney et al. 2014b; Thudi et al. 2014; Saxena et al. 2014; Bajaj et al. 2016; Gowda et al. 2011).

In addition to hormonal control systems regulating branch outgrowth including auxin, cytokinin and strigolactone (Beveridge and Kyozuka 2010; Ongaro and Leyser 2008), branching is also known to be influenced by the flowering control system. Flowering time has been associated with growth habit and branching in several species including legumes (Yang et al. 2017; Beveridge et al. 2003), and specifically in chickpea (Bonfil et al. 2006b; Aryamanesh et al. 2010). This suggests some degree in overlap in the genetic control of these traits, either through pleiotropic effects of single genes or close linkage of different genes. One well-known illustration of this type of pleiotropy is the case of genes in the *FT* family, which were initially identified as floral integrators but subsequently shown to be involved in a wider range of developmental processes (Pin and Nilsson 2012).

Chapter 4 investigated the genetic control of flowering in chickpea using several inter- and intraspecific populations. A cluster of *FT* genes located on chromosome 3 were identified as the most likely candidates to explain the flowering time difference observed in the interspecific crosses, and could potentially be key target genes in the domestication process of chickpea. In the process of that study, it became apparent that differences in flowering were accompanied by substantial variation in plant architecture. This chapter explores the genetic control of branching and growth habit in the interspecific CRIL2 population (CRIL2) and its relationship to the major flowering time QTL, QTL3a, and to the *FT* gene cluster.

5.2 Materials and methods

Plant material and growing conditions

In this study we used an interspecific population (CRIL2) comprising 128 recombinant inbred lines (RILs) derived from the cross between the *desi* cultivar ICC4958 (*C. arietinum*) and the wild accession PI489777 (*C. reticulatum*). Further information about the parents as well as growing conditions used in this chapter are described in Chapter 4.

Phenotypic evaluation and genetic analysis

Two different phenotypic measures were used to assess overall shoot architecture; branching and growth habit. Branching tendency was quantified by recording the length of the main shoot and of every branch in all plants three weeks after emergence. The ratio of total branch length to main shoot length (branching index, abbreviated as BI) was calculated to normalize for differences in general vigour. Growth habit was scored using a four-category scale, according to a visual assessment of the angle of the branches from the vertical axis at harvest stage, and each category was assigned a value from 1 to 4, as follows: 1 - prostrate (branches parallel to the ground, $0\text{-}10^0$), 2 - semi-prostrate (branches with $10\text{-}45^0$), 3 - semi-erect (branches between $45\text{-}70^0$) and 4 - erect (branches $> 70^0$). The mean value from 4 replicate plants (replications) for each RIL and parent line was used for analysis. QTL analysis was performed as described in Chapter 4, using the same genetic map.

5.3 Results

5.3.1 Phenotypic evaluation

Table 5.1 summarizes the mean values obtained for the three traits analysed (flowering time, branching index and growth habit) in the different conditions. Although flowering time was thoroughly addressed in Chapter 4, a summary of the mean values for DTF was included in this table because the relation between flowering time and GH or BI will be evaluated in this chapter.

Table 5.1 Comparison of mean and range values obtained for 3 traits analysed in CRIL2 parents and RILs grown under long and short photoperiod.

	Days to flowering		**Branching index**		**Growth habit**
	Long days	Short days	Long days	Short days	Short days
RILs					
Mean ± sd	29.8 ± 2.0	101.5 ± 33.4	0.45 ± 0.3	0.7 ± 0.4	2.5 ± 1.1
Range	23 - 35	45 – 130[a]	0 - 1.37	0.07 – 1.57	1-4
ICC4958					
Mean ± sd	29.5 ± 2.9	60.75 ± 1.0	0.42 ± 0.3	1.04 ± 0.4	4 ± 0.0
Range	26 – 33	60 – 62	0 – 0.64	0.39 – 1.32	4
PI489777					
Mean ± sd	33.2	130[a]	0.28 ± 0.2	0.77 ± 0.3	1 ± 0.0
Range	30 – 37	-	0.13 – 0.57	0.49 – 1.19	1

[a] Plants unable to flower by 11-04-2016 were scored with a DTF value of 130 (days from emergence to end of scoring period)

5.3.1.1 Branching ratio

Photoperiod is a well-known environmental factor that influences bud formation and dormancy in many species (Leduc et al. 2014). Figure 5.1 shows that the domesticated parent ICC4958 displayed a similar architecture under both photoperiods. No major alteration in the number of primary or secondary branches was found, and the only difference apparent was the length of the primary branches, which were longer in SD than in LD. By contrast, shoot architecture of the *C. reticulatum* accession PI489777 was completely different in the two conditions. Under LD, it resembled ICC4958 with a dominant main shoot and relatively few primary branches (Fig 5.1 B). In contrast, under SD, apical dominance was lost in PI489777 to the point that the main stem

was not identifiable, and the number of both primary and secondary branches was increased resulting in plant with a rosette-like phenotype (Fig 5.1 A).

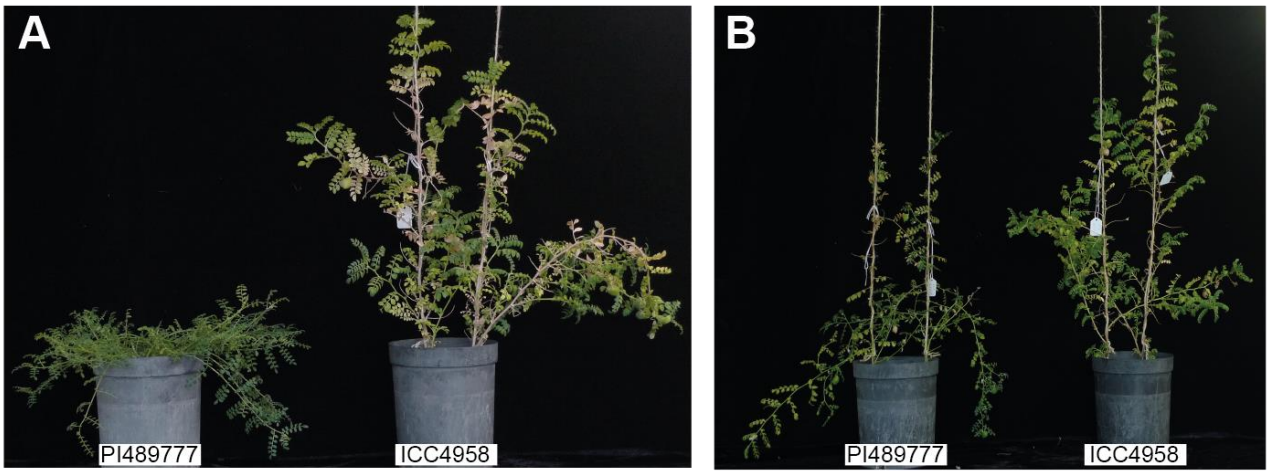


Figure 5.1 Growth habit phenotype of CRIL2 parental lines grown under SD (A) and LD (B) photoperiods.

As described in Chapter 4, these lines also differed dramatically in flowering time under SD, raising the possibility that the difference in architecture might be a consequence of the difference in flowering time. In an attempt to exclude this possibility, branching index was calculated in plants at a relatively early stage of development (three weeks old) where flower buds were not yet visible in either line, and the major differences in architecture shown in Figure 5.1 were not yet obvious. In ICC4958, BI showed no statistical difference between photoperiods ($t(6) = -2.357$, $p = .057$), while BI of wild chickpea accession PI489777 was significantly lower in LD than in SD ($t(6) = -2.656$, $p = .038$). This result indicates that the difference in branching behaviour of these accessions, is already present in plants as young as three weeks, suggesting that it is not simply a consequence of their difference in flowering status. BI was also found to be significantly lower in LD than in SD for the RIL population overall ($t(124) = -11.786$, $p = 5.67E-022$). Also noteworthy, the mean BR of the RILs was similar to that of cultivated parent under LD and to that of wild parent under SD, suggesting that the differential response above described for PI489777 is maintained in the population. Interestingly, transgressive segregants with lower or higher BI than parental lines were evident, suggesting genetic interaction between alleles present in the parental lines.

5.3.1.2 Growth habit

Photoperiod had also a major impact on the growth habit of CRIL2. Under the 16-hour LD conditions used in this experiment, the phenotypic differences in plant architecture (including growth habit) were subtle even in the parental lines (as previously shown in figure 5.1), and difficult to assess, so we therefore decided to assess growth habit only under SD, where there was wide variation.

Under SD, the differences in growth habit of both parents were clear; ICC4958 displayed the expected erect phenotype whereas branches on PI489777 were almost parallel to the ground. Among the RILs a wide range of phenotypes were found, covering the whole spectrum between the parental phenotypes. This is illustrated in Figure 5.2. Out of 124 RILs, 32 were classified as prostrate, 39 semiprostrate, 23 semierect and 30 erect (Figure 5.2).

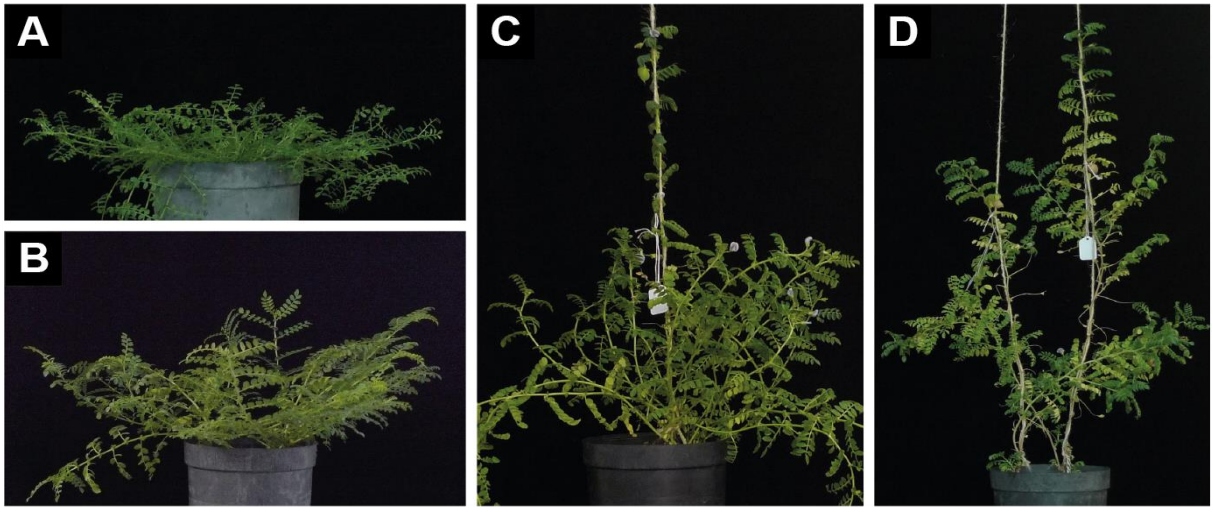


Figure 5.2 Representative plants illustrating the four growth habit categories used for classification of the CRIL2 population grown under SD condition: prostrate (**A**), semiprostrate (**B**), semierect (**C**) and erect (**D**).

5.3.1.3 Correlation between flowering, branching and growth habit

General observations suggested that late flowering lines in the population also showed a growth habit and branching pattern that resemble the wild profile, so we investigated the correlation between flowering (SD and LD), growth habit (SD) and branching ratio (SD and LD). Spearman's rank coefficient values obtained for each pair of traits showing significant

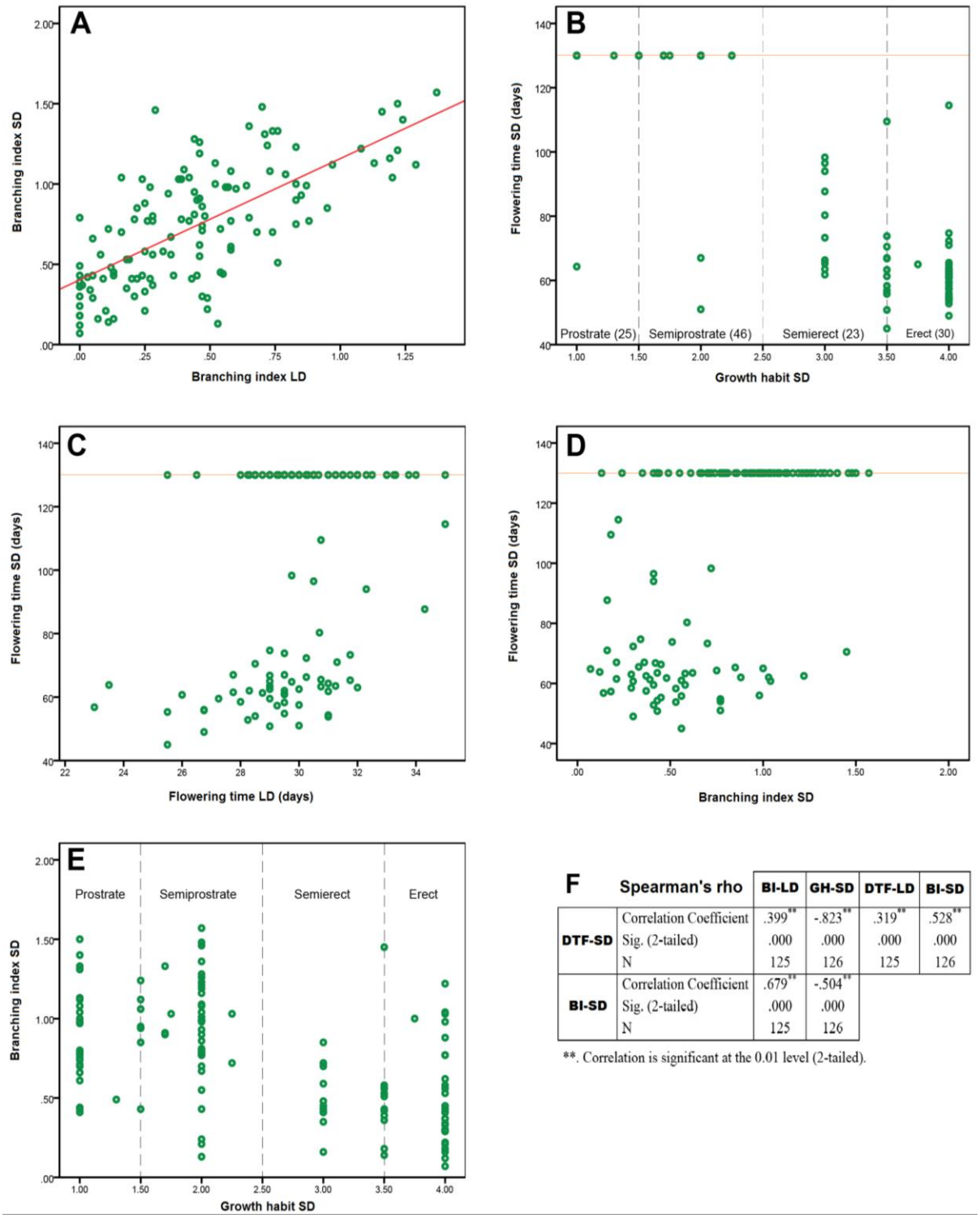
correlation can be found in Fig 5.3 (F). The correlation between DTF under both photoperiods (Fig 5.3 C) was evaluated in chapter 4 and no more information will be added in this section.

A very strong negative correlation (r = -.823, p< 0.001) was found between flowering time and growth habit under SD (Fig 5.3 B), confirming that in the segregating population, prostrate growth habit is associated with late flowering, as expected. Inspection of individual RILs showed that this correlation was nearly perfect, with flowering observed in all 53 RILs with an erect or semierect phenotype but in only 3 out of 71 lines categorized as prostrate or semiprostrate.

A strong negative correlation (r = -.504, p< 0.001) was also found between growth habit and branching ratio (Fig 5.3 E), indicating that erect and semierect plants in general also had a lower branching index.

Moderate correlation (r = .399, p< 0.001) was found between flowering time and branching ratio in SD (Fig 5.3 D), and no correlation between these traits was found in LD (r = -.014, p= .875); in this condition the mean DTF of the RILs was similar to that of the early parent (table 5.1), shortening the vegetative growth period.

A significant, strong, positive correlation (r = .679, p< 0.001) was found between BI in LD and in SD (Fig plot A). This indicates that the genetic differences influencing branching prior to flowering are expressed under both photoperiods.



F

Spearman's rho		BI-LD	GH-SD	DTF-LD	BI-SD
DTF-SD	Correlation Coefficient	.399**	-.823**	.319**	.528**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	125	126	125	126
BI-SD	Correlation Coefficient	.679**	-.504**		
	Sig. (2-tailed)	.000	.000		
	N	125	126		

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 5.3 Relationship between branching index (BI), growth habit (GH) and flowering time (DTF) under long day (LD) and short day (SD) photoperiod (**A** to **E**). Spearman correlation coefficient value obtained for each pair of traits (**F**). Numbers in parentheses in (B) represent the total number of RILs in each category. Plants not able to flower under SD were scored with a DTF of 130 (threshold indicated by orange line in B, C and D), coinciding with the end of the scoring period.

5.3.1 QTL analysis

As shown in Table 5.2 and Figure 5.4, QTL analysis identified 3 significant QTLs controlling growth habit and branching ratio, on two different chromosomes. For growth habit, one major QTL (QTL_{GH3}) spanning 1.3 cM was found on LG3. This QTL explained 66% of the phenotypic variation for growth habit, and the maximum LOD score (34.03) was obtained for the *FTa1* marker. A minor QTL_{GH4} (5.9 % PEV) was also identified for this trait in LG4, spanning 9.2 cM and with marker S360p1277380 showing the highest LOD (5.56).

Table 5.2 Summary of the QTLs found in CRIL2 for growth habit and branching index, indicating the LOD value and the proportion of phenotypic variance explained (PEV).

Trait	Condition	LG	QTL	Marker[a]	LOD	PEV	Threshold[b]
Growth habit	Short days	LG3	QTL_{GH3}	*FTa1*	34.03	66.6	2.6
		LG4	QTL_{GH4}	*S360p1277380*	5.56	5.9	3.1
Branching		LG3	QTL_{BI3}	*FTa1*	10.64	33.1	2.65
		LG6	QTL_{BI6}	*S826p761274*	2.69	6.5	2.85
Branching	Long days	LG3	QTL_{BI3}	*FTa1*	5.45	18.4	2.5

a. Marker with the highest LOD score
b. LOD score necessary for a 99.5 confidence, obtained by permutation test for each trait and linkage group

One significant major QTL_{BI3} was also found for branching index also on LG3 between markers SUVH4 and WUS11/CDF2d, and it was consistent across both photoperiods. This QTL displayed a higher LOD (10.64) and proportion of phenotypic variation explained (33 %) in SD than in LD (LOD = 5.45; 18 % PEV), suggesting that the effect of the gene underlying this QTL is more prominent under short photoperiods.

Interestingly, QTL_{GH3} and QTL_{BI3} are delimited by the same markers, and are therefore both located in the same region as the flowering time QTL_{3A} (Chapter 4). Also, for both QTL, *FTa1* was again the marker with highest LOD score (Fig 5.4). These similarities strongly suggest that the differences in flowering time, branching index and growth habit are likely to be pleiotropic effects of the same gene, although the alternative possibility of independent control by tightly linked genes cannot be excluded.

Although there was no other significant QTL detected in the analysis, one other potential region in LG6 (tentatively termed QTL_{BI6}) had a LOD score very close to the threshold indicated by permutation test (2.85), and explained 6.5 % of the variation.

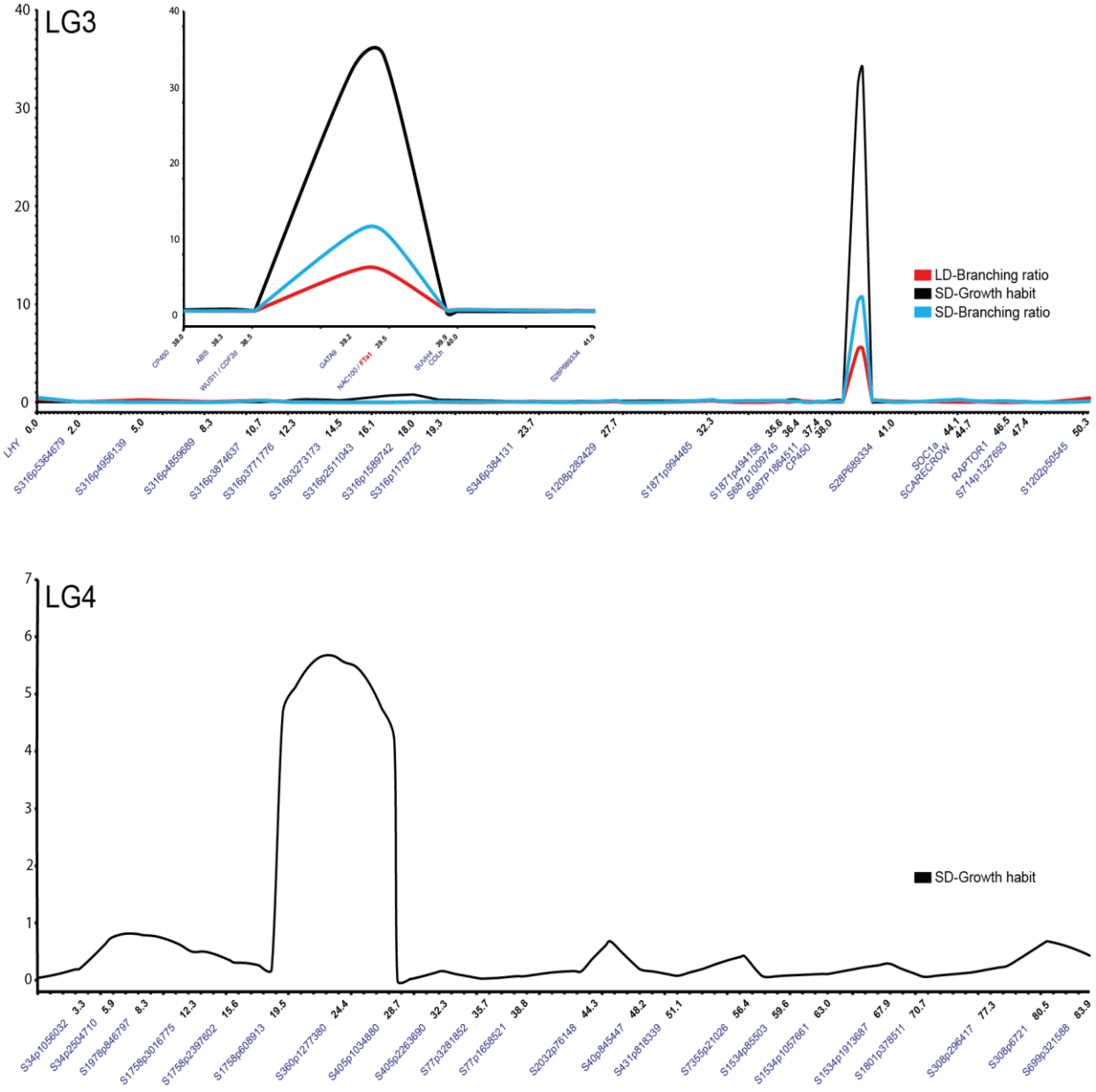


Figure 5.4 QTLs detected for branching ratio and growth habit in the interspecific population CRIL2. The horizontal axis represent linkage groups 3 (up) and 4 (down), with markers indicated in blue and distances (in cM) in black. Vertical axis indicates LOD score.

5.4 Discussion.

Shoot architecture is defined as the three-dimensional organisation of the parts of the plant above the ground. It is a very broad term that includes the size, shape, position and angle of branches, leaves and floral organs (Reinhardt and Kuhlemeier 2002). Shoot architecture is one of the main determinants of yield in crops generally (Huyghe 1998) and in chickpea, its influence on productivity has also been documented (Siddique and Sedgley 1985; Rubio et al. 2004). Modification of this trait may therefore provide potential to boost the globally low productivity of chickpea (Millan et al. 2006; Varshney et al. 2014a), but the nature of its genetic and physiological control has so far received little attention. Since overall shoot architecture is determined both by growth habit and the degree of branching, a better understanding of both traits will be important in order to improve it and to assist with managing it in breeding programs.

In this study, we present 3 significant QTL for branching index (one QTL in LG3) and growth habit (one QTL in LG3 and another in LG4). Both QTLs obtained here on chromosome 3 were defined by the same marker interval. Interestingly, they also perfectly co-located with the major QTL reported for flowering time on this population (described on chapter 4), between markers CDF2d and SUVH4. The possible causes of the co-location of the QTL for these traits will be discussed later, but it is important to note that co-location of flowering and growth habit loci in this region of chromosome 3 has previously been reported in another chickpea interspecific cross (Aryamanesh et al. 2010), and a similar relation of flowering with different aspects of aerial architecture have been previously described in other legume species and chickpea itself; Lichtenzveig et al. (2006) found that all loci associated with chickpea flowering time also had a significant effect on the number of branches. Similarly, the soybean major flowering-locus *E1* has also been associated with variation in branch number (Yang et al. 2017). In *Medicago truncatula*, QTLs for flowering time and several branching-related traits (including number of primary and secondary branches, length of main stem and branch elongation rate) were found to be colocated in a region of chromosome 7 that includes the orthologous *FT* cluster described forchickea in Chapters 3 and 4(Julier et al. 2007; Lagunes Espinoza et al. 2012).

Branching

Previous studies on branching in chickpea are focused mainly in the number of branches, so the BI used in this study is a novel approach to study the genetic of this trait. One of the potential complications with an assessment of branching tendency is the influence of the reproductive state of the plant in this trait. In this study, branching was quantified at a relatively early stage of development in order to avoid an indirect effect of differences in flowering time. Interestingly, the major locus controlling this trait in both SD and LD (QTL_{BI3}) mapped to a 1.4 Mbp region of chromosome 3 between markers SUVH4 and CDF2d, in which a major locus controlling SD flowering was also identified (Chapter 4).

Chickpea chromosome 3 has been previously linked with branching-related traits in a number of different studies. In an interspecific cross also using the same parents, Saxena et al. (2014) identified 8 QTLs for number of branches in 5 different LGs, including one in LG3 between markers CaSSR106 and CaTFSNP87. CaSSR106 is located at 4.7 Mbp, in a position very close to the top of the chromosome, which together with the close linkage of CaTFSNP87 (2.8 cM apart) indicates that is a different QTL to the one identified here. Gowda et al. (2011) described three QTLs for number of branches in LG3, and two of them occupies a central position in the LG, wich suggest a co-location with QTL_{BI3}. Unfortunately, the lack of common markers makes it impossible to determine their exact relationship. Interestingly, QTLs for number of branches and plant height (which can be partly a manifestation of GH) also co-locate in Gowda et al. (2011), similar to our results. Bajaj et al. (2016) described seven genomic regions explaining a major proportion of the variation for branch number in domesticated chickpea using genome-wide association analysis. One of these was an auxin transporter-like (NCBI gene ID 101505743) located in chromosome 3 but in a position distant (19.2 Mbp) and quite distinct from the branching QTL presented here, suggesting that this may not be the same locus.

Growth habit

Two QTLs were identified for growth habit in CRIL2. One of them, a minor QTL, was mapped in chromosome 4, while the second (QTL_{GH3}) is major QTL placed in chromosome 3, defined by the same interval than the QTL above described controlling branching ratio. Two major loci have been reported to date controlling growth habit in chickpea. All reports using interspecific

populations converge over the same region of chromosome 3, pointing to the existence of a locus responsible for the major part of the differences in growh habit between *C. ariteinum* and *C. reticulatum*. This locus, called *Hg*, was first described using interspecific crosses between the wild accession PI489777 and the cultivated lines PI374080 and PI360168 (Muehlbauer and Singh 1987; Kazan et al. 1993). *Hg* was initially linkedin these studies to the isozyme locus 6-phosphogluconate dehydrogenase, and later was mapped to chromosome 3 between the markers STMS10 and TA64 by Winter et al. (2000) using a population derived from the same parents as CRIL2. Similarly, Cobos et al. (2009) mapped a gene controlling interspecific growth habit differences between TA6 and TA64 markers in LG3. Ali et al. (2015) found significant association of some markers in LG3 (TA142 and STMS5) with GH, although it was not detected by QTL analysis. Since QTL_{GH3} perfectly co-locates within the intervals reported in these three studies, it is likely that they all refers to *Hg*. Finally, Aryamanesh et al. (2010) found a linkage between growth habit and the marker TA142, which is physically very close to *CaFTa1* (<150kb away).

A second gene (*Hg2*) was recently described in LG1 affecting growth habit in both inter and intraspecific populations (Ali et al. 2015). However, in the present study, we found no association of growth habit with this chromosome. A wide range of variation for growth habit has been described in cultivated chickpea germplasm (Upadhyaya et al. 2006), so other loci in addition to *Hg* and *Hg2* are likely to participate in control of this trait. A recent genome-wide association study also found significant association with markers on chromosomes 4, 6 and 7, and identified five candidate genes that were validated using intra- and interspecific crosses and differential expression (Upadhyaya et al. 2017).

Co-location of QTLs for growth habit, branching ratio and flowering time

As mentioned above, QTLs for growth habit, brainching index and flowering show very close co-location within the same region of chromosome 3. Two different scenarios could explain this co-segregation: (1) the existence of different but tightly linked genes affecting different traits or (2) the pleiotropic action of a single gene. In the first case, we need to consider what is known about the regulation of shoot branching in other species in order to propose candidates with potential to regulate these traits. Genes related with the metabolism of hormones (especially auxins) would deserve special attention, as they are the main factors regulating shoot branching,

as documented by many reviews (Janssen et al. 2014; Rameau et al. 2015; Domagalska and Leyser 2011; Kebrom et al. 2013; Wang and Li 2008; Teichmann and Muhr 2015). The flow of auxin in the main stem is responsible for maintenance of apical dominance and the repressing the outgrowth of new branches. Auxin also modifies the circulating levels of two other main hormones involved in bud outgrowth; it promotes upregulation of strigolactones (also bud growth inhibitors) and at the same time reduces the synthesis of cytokinins, which are branch inducers (Ongaro and Leyser 2008; Beveridge and Kyozuka 2010; Domagalska and Leyser 2011). In addition to the hormonal control, local regulation further contributes to the control of bud activation. Little is known about this aspect, and the only well described regulators are the Arabidopsis transcription factor *BRANCHED1* (*BRC1*) and *BRC2* and their homologs in other species; *TEOSINTE BRANCHED1* in maize and sorghum, *SlBRC1a* and *b* in tomato, *PsBRC1* in pea and *FINE CULM 1* in rice (Minakuchi et al. 2010; Braun et al. 2012; Martin-Trillo et al. 2011; Finlayson 2007; Takeda et al. 2003; Aguilar-Martínez et al. 2007; Doebley et al. 1997). *BRC1* is specifically expressed in axillary buds and it has been proposed that these genes integrate endogenous signals from both strigolactones and cytokinins with environmental cues such as light or nutritional plant status (Aguilar-Martínez et al. 2007; Teichmann and Muhr 2015; Hubbard et al. 2002; Poza-Carrion et al. 2007).

The control of the angle adopted by lateral branches is a phenomenon that only recently is becoming understood but seems to be a trait mainly determined by gravitropism. In the main stem, deviations from vertical lead to the creation of an asymmetric auxin distribution that causes differential elongation on two sides of the stem and a return to vertical growth (Morita 2010; Roychoudhry and Kepinski 2015). In the lateral branches, this gravitropic response is opposed by an antigravitropic offset (AGO). The magnitude of this AGO is controlled by auxins in the statocytes (gravity-sensing cells) of the lateral shoots and determines the growth angle of the shoot branches; a weaker AGO enables vertical growth and vice versa (Roychoudhry et al. 2013). Auxin, therefore, plays also a central role in this trait, but several other genes have been shown to regulate lateral branching in different plant species (Roychoudhry and Kepinski 2015).

Although the presence of more than one gene affecting each trait independently can not be discarded, the second scenario, involving a single gene with pleiotropic effects, is a simpler and more likely option, in view of the similar effect of photoperiod on phenotypic expression of the

three QTLs (DTF, GH and BR). The *CaFTa1-FTa2-FTc* cluster was proposed as the most likely candidate for the flowering QTL (see chapter 4) and a *CaFTa1* marker also was the most tightly associated with growth habit and branching ratio, so their suitability as candidates to explain also the phenotypic differences in growth habit or branching will be discussed.

The pleiotropic effect can be exerted in a direct or indirect manner. There is no clear previous precedent for how FT proteins could directly influence the three traits, but genes in this family are a good example of great functional diversification. They were first discovered through their role in flowering (Kobayashi et al. 1999), but since then *FT* homologs have been involved in a much wider range of developmental processes such as stomatal control, bud set and photoperiodic growth control, inflorescence meristem identity, bulb formation and tuberization (Pin and Nilsson 2012). Importantly, *FT* genes are also considered to be general regulators of plant architecture, through effects on lateral branching and meristem determinacy, not only in Arabidopsis (Hiraoka et al. 2013; Huang et al. 2013) but also in tomato (Lifschitz et al. 2006; Weng et al. 2016), rose (Randoux et al. 2014) and rice (Tamaki et al. 2007; Tsuji et al. 2015). In *Brassica napus*, a particular *FT* allele was found to affect both flowering and plant height (Schiessl et al. 2015; Shi et al. 2009). In pea, the late-flowering phenotype of some photoperiod response mutants with misregulation of *FT* genes is coupled with a profusion of basal branches (Weller et al. 1997a; Ridge et al. 2016; Hecht et al. 2007b). In Medicago, *fta1* mutants show a highly-branched, prostrate phenotype under LD similar to that seen in wild chickpea under SD, while the overexpression of *MtFTa1* has the opposite effect and confers a more erect phenotype under SD (Laurie et al. 2011).

While little is so far known about the mechanisms linking *FT* genes to these processes, especially in legumes, this question has been explored to a limited extent in Arabidopsis. A model has been proposed in which the branch-inhibiting transcription factor *BRC1* interacts with Arabidopsis *FT* in axillary buds and this results in mutual inactivation such that when the complex *FT*/*BRC1* is formed, *FT* cannot promote flowering but *BRC1* is also inactive, so the branching is possible (Rameau et al. 2015). In legumes, the phenotype of pea *brc1* mutant have been characterized and is consequent with this idea, as they show a higher number of branches with no alteration of flowering time (Braun et al. 2012). As above mentioned, *BRC1* plays a role in branching control similar to that of *FT* in flowering, integrating signals from multiple pathways, including

strigolactones and cytokinins. Therefore, this gene would be the common link between the hormones and *FT* genes in the control of branching (Rameau et al. 2015).

All these reports above presented support the case for a direct connection between *FT* gene(s) and the pathways regulating shoot architecture that could explain the colocation of the three QTLs obtained in chickpea in this study. However, a pleiotropic effect of this nature does not necessarily mean that *FT* genes participate directly in the regulation of these traits, as they may also be an indirect consequence of flowering. *FT* floral induction implies the regulation of genes belonging to the MADS-box family, which are necessary for the vegetative to inflorescence meristem transition and flower development. In addition, MADS-box genes participate in many other plant processes, some of them key for plant architecture such as root, stem and leaf development (Becker and Theißen 2003; Pelaz et al. 2003). It is therefore plausible that some the genes regulated by the downstream signalling cascade initiated by *FTs* also participate in the molecular pathways controlling branch number and/or angle. Whether through the action of MADS-box genes or other unknown factors, it is clear that the transition to the reproductive stage is accompanied by many other changes in allocation of resources, and in many species this includes changes in shoot architecture such as increased apical dominance and decreased lateral branching (Davies 1995; Rameau et al. 2015; Banfield and Brady 2000).

Conclusion

In summary, the region of chickpea chromosome 3 between markers SUVH4 and CDF2d is associated not only with flowering but also with differences in growth habit and branching between domesticated chickpea and *C. reticulatum*. Further research is necessary in order to clarify whether there may be several different genes in this region affecting different traits or whether these differences reflect the pleiotropic effects of a single gene/locus, which seems likely to be the *FT* cluster. In either case, a better understanding of the molecular basis for this association may help identify strategies to control them independently of each other.

Chapter 6. Exploring natural sequence variation within the *FTa1-a2-c* cluster

6.1 Introduction

Chapter 4 of this thesis presented evidence supporting a potential role for the *FTa1-FTa2-FTc* cluster as the most likely candidate for a major QTL on chickpea chromosome 3 that determines flowering time differences between wild and domesticated chickpea. In the case that one or more of these genes are responsible for the QTL effect, the early (domesticated) allele would be most likely to carry a gain-of-function mutation, considering that these genes in other temperate legumes are known to promote flowering. Consistent with this hypothesis, we showed an upregulation and early induction of these genes in the early parents of the three crosses used to define the QTL. Although it is possible that a gain-of function mutation can result from a change in the coding region, such mutations are more likely to arise in the regulatory regions, leading to upregulation of genes.

For flowering time in particular, there are several examples of how allelic variation in the non-coding sequences of key genes can make a major contribution to phenotypic variation and geoclimatical adaptation. The best-known example is *FLOWERING LOCUS C* (*FLC*) in *Arabidopsis thaliana*, an important gene that is a key determinant of the vernalization response. Sequence variation at non-coding sites of the *FLC* locus is a major factor contributing to differential flowering responses to cold and associated ecogeographic adaptation (Li et al. 2014). Similar to Arabidopsis, variation in heading date of *Brassica oleracea* has also been achieved through selection of sequence variants in UTR sequences in an *FLC* ortholog (Irwin et al. 2016). The regulation of *FLOWERING LOCUS T* (*FT*) has also been well studied. In Arabidopsis, variation in the promoter sequence has been associated with differential *FT* expression and phenotypic variation (Schwartz et al. 2009). Changes in the length of *FT* promoter due to various insertions and deletions confers different responsiveness to photoperiod, playing an important role in the adaptation of this species to different latitudes (Liu et al. 2014). In sunflower, a frameshift mutation in a *FT* paralog also plays a key role in domestication, experiencing a selective sweep during the process (Blackman et al. 2010).

The *FTa1-FTa2-FTc* cluster is conserved throughout the temperate galegoid legumes, and its importance has previously been explored in pea and Medicago. In both species, the *FTa1* and

FTc genes have been shown to be effective floral inducers, and the *FTa1* gene in particular has been shown to have a key role in induction of flowering and mediation of the response to vernalization (Laurie et al. 2011; Putterill et al. 2013; Hecht et al. 2011). However, despite the importance of *FTa1* and other *FT* genes in legumes, little is known about the regions of these genes that are necessary for their characteristic patterns of regulation (Jaudal et al. 2013).

With all this in mind, this chapter aimed to investigate the variability in the sequence of the *FT* cluster within a collection of wild and cultivated chickpea accessions. The detection of the polymorphisms at both the inter- and intraspecific level is a first, necessary step to identify regions that may modulate *FT* expression in chickpea. The variants thus identified can be targeted in future studies to evaluate their functional and adaptive significance, not only in chickpea but likely in other phylogenetically related legumes.

6.2 Materials and methods

Plant material and DNA extraction

A total of 96 chickpea lines were selected for analysis (Appendix 6.1). These included the six parent lines (two *reticulatum*, four *arietinum*) of the populations described in Chapter 4, and three additional *C. arietinum* lines that have been used as parents in other populations for which flowering time QTLs have been reported (Table 6.1).

Table 6.1 Chickpea accessions used as parental lines in crosses with described phenological QTLs and the references of the studies reporting them. Accession used as parents in the three crosses described in chapter 4 of the present dissertation are indicated.

Accession	Reference
ICCV2	(Pushpavalli et al. 2015; Cho et al. 2002; Vadez et al. 2012; Jamalabadi et al. 2013)
Cr5-9; ICCL81001	Parental line of RIP12. (Cobos et al. 2009)
CA2156	(Cobos et al. 2007)
ILC3279	Parental line of RIP5 and RIP8. (Ezzat et al. 2015) (Rehman et al. 2011; Jamalabadi et al. 2013; Hamwieh et al. 2013a)
ICC4958	Parental line of CRIL2. (Varshney et al. 2014a; Samineni et al. 2016; Das et al. 2015b; Gowda et al. 2011)
WR315	Parental line of RIP5 and RIP8. Chapter 4 of the present dissertation
JG62	(Cho et al. 2002; Cobos et al. 2007; Vadez et al. 2012) (Gowda et al. 2011)
PI489777	Parental line of CRIL2. (Samineni et al. 2016)

The remaining 87 chickpea accessions were selected from a larger core collection of 256 lines, kindly supplied by Dr Tim Sutton from the South Australia Research and Development Institute (SARDI). The selection aimed to maximize the variation in flowering time and geographical origin, according to flowering data from field trials at Turretfield Research Centre in South Australia (see Electronic Supplementary Material 1) also kindly supplied by Dr Sutton.

Growing conditions and flowering time phenotyping

Three plants of each of the selected lines were grown in the phytotron at the University of Tasmania (Hobart) in two consecutive years; during 2015, the 87 lines belonging to the larger collection were grown in long day conditions (LD), whereas in 2016 the same 87 accessions plus the 9 parental lines were grown under both LD and short day (SD) photoperiods, as described in chapter 2. Three plants of each accession were sown in each of the environments and seasons.

The time in days from emergence to first open flower (DTF) was recorded for each plant and the mean value for each line was used for further data analysis.

PCR amplification

To sequence the approximately 57 kb region containing *FTa1*, *FTa2* and *FTc*, six overlapping fragments were amplified separately using the primers described in Table 6.2. All PCRs were performed in a final volume of 30 µL containing 150 ng of template DNA, 0.4 µM of each primer and 15 µL Ranger Mix (Bioline Australia Pty. Ltd). Reactions were conducted in a thermal cycler with heated lid, with the following program: an initial denaturation of 2 minutes at 95 ºC and 35 cycles of 98 ºC for 10 seconds and an annealing/extension step as specified in table 6.2.

Table 6.2 Primers, annealing temperature and extension time used to amplify *FTa1*, *FTa2* and *FTc* in 96 chickpea lines.

	Primer name	Primer sequence	Annealing TºC / Extension time
Fragment 1			
	CaFTS-F2	TGAAACGGCCAAAGTTACGG	58 ºC 11 min
	CaFTa1-R1	CAAAACGATGAATTCCAGAGG	
Fragment 2			
Normal lines	CaFTa1-F4	TGCCGGTGAATTATATGGGCCG	58 ºC 10 min
	CaFTa1-R4	GTGGGGCGACCTTTGTTTGC	
Deletion lines	CaFTa1-F4	TGCCGGTGAATTATATGGGCCG	
	CaRMK-R3	CAATGCTATTAGTTACTACGTCGAC	
Fragment 3			
Normal lines	CaFTa2-F6	AAGCCCACAACCCACCTAAGGG	58 ºC 10 min
	CaFTa2-R1b	ACTAGCCCCAGCAGTTGAAG	
Deletion lines	CaRMK-F1	ACTGTTCTGCACACAGTGGCTACC	
	CaFTS-R7	AGGCCAAAGACAAGATCCCG	
Fragment 4			
	CaFTa2-F1	TAGGCGGAAACGATCTCAGG	58 ºC 10 min
	CaFTa2-R6	GCCATAAACCTCTGTCGAACGGC	
Fragment 5			
	CaFTc-F4	AACCGTTCTCCACACTCAGC	60 ºC 11 min
	CaFTc-R5	CAGGGGGAGTTGTGTCTTGG	
Fragment 6			
	CaFTc-F2b	ACCCACCCAACAAATCTCCC	60 ºC 11 min
	CaFTc-R1b	TTCTCGGCAATCGTAGGTCG	

PCR product purification and DNA quantification

PCR products were purified using ISOLATE II HT 96 Clean-Up Kit (Bioline Australia Pty. Ltd) following supplier's instruction. Accurate quantification of the purified PCR products was obtained using a Standard Sensitivity Large Fragment Analysis kit in a Fragment Analyzer (Advanced Analytical Technologies, Inc) according to the manufacturers' instructions.

Library preparation and sequencing.

Paired-end libraries of ~300-400 bp were prepared and indexed for all samples and normalized using a Nextera XT Library Preparation Kit (Illumina Inc.). The appropriate size and concentration of the libraries was confirmed using a Standard Sensitivity Large Fragment Analysis kit in a Fragment Analyzer (Advanced Analytical Technologies, Inc) following manufacturer instructions. Sequencing of the libraries was performed using a 300-cycle MiSeq Reagent kit v2 in a MiSeq system (Illumina Inc).

Sequence data processing

Raw sequence data was processed as follows. Sequences without index were eliminated and the remaining sequences were sorted into their corresponding sample based on their index. The quality of the reads was then evaluated with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and sequences with a Phred quality score (Q score) of less than 30 were eliminated.

De novo assembly and mapping to reference

Two different strategies were used to ensure the accuracy of the final sequences. First, a *de novo* assembly was performed for each sample using Geneious 8 and the following options: maximum gaps per read = 10%; maximum gap size = 1; word length = 24 and Index word length = 14; maximum mismatches per read =4; maximum ambiguity =4. Words repeated more than 100 times were ignored and the distance between paired reads was used to improve assembly, with an expected value of 0.

Second, the paired-end reads were mapped against the *FT* cluster from the reference genome of cultivar CDC frontier available in NCBI, and a consensus sequence was obtained for each

accession. Mapping was performed using Geneious 8 mapper and the following parameters: maximum gaps per read = 10%; maximum gap size = 6; word length and Index word length = 10; maximum mismatches per read =4; maximum ambiguity =4. A distance between paired reads of 0 bases was used as indicator to improve mapping.

The scaffolds generated by *de novo* assembly and the consensus obtained from mapping to the reference were aligned and compared. When no differences were observed between these sequences, the consensus was considered as the final sequence. In the case of discrepancies, these were inspected in both the mapping and assembly. The consensus was then manually edited to reflect the correct sequence. To test the accuracy of the corrections made in this way, a new mapping using the improved consensus as reference was performed and visually inspected to verify the quality of the changes.

Cloning of PCR products and PCR colony

The cloning of the PCR products and PCR colony were realized following the procedure described in Chapter 2 (See sections 2.5 and 2.4.2).

Statistical analysis

The software DnaSP v5.10.01 (Librado and Rozas 2009) was used to explore nucleotide diversity. Two parameters (π and θ) were calculated, where π is the average number of nucleotide differences per site between any two DNA sequences and θ is derived from the total number of segregating sites and related to sample size. Indel diversity (π_i) as well as the number and diversity of the haplotypes were also calculated using the same software.

The frequency distributions, analysis of the correlation between flowering time in different environments and the statistical significance of some detected polymorphisms in the flowering phenotype (tested through independent sample t-test) were performed using SPSS software (IBM).

Phylogenetic analysis

For phylogenetic analysis, a neighbour-joining tree was constructed in Geneious 8 software using the Tamura-Nei genetic distance model and bootstrap resampling method with 1000

replications. The Median-joining network (Bandelt et al. 1999) with an epsilon value of 0 was calculated using PopART (http://popart.otago.ac.nz).

6.3 Results

6.3.1 Flowering phenotype of the chickpea collection

A chickpea selection of 96 lines with diverse geographical origin (Fig 6.1) including *desi* and *kabuli* as well as two *C. reticulatum* accessions was grown in controlled environment facilities at University of Tasmania (Hobart) under two different photoperiods (LD and SD). The same population was previously grown in field conditions at Turretfield research station (South Australia).

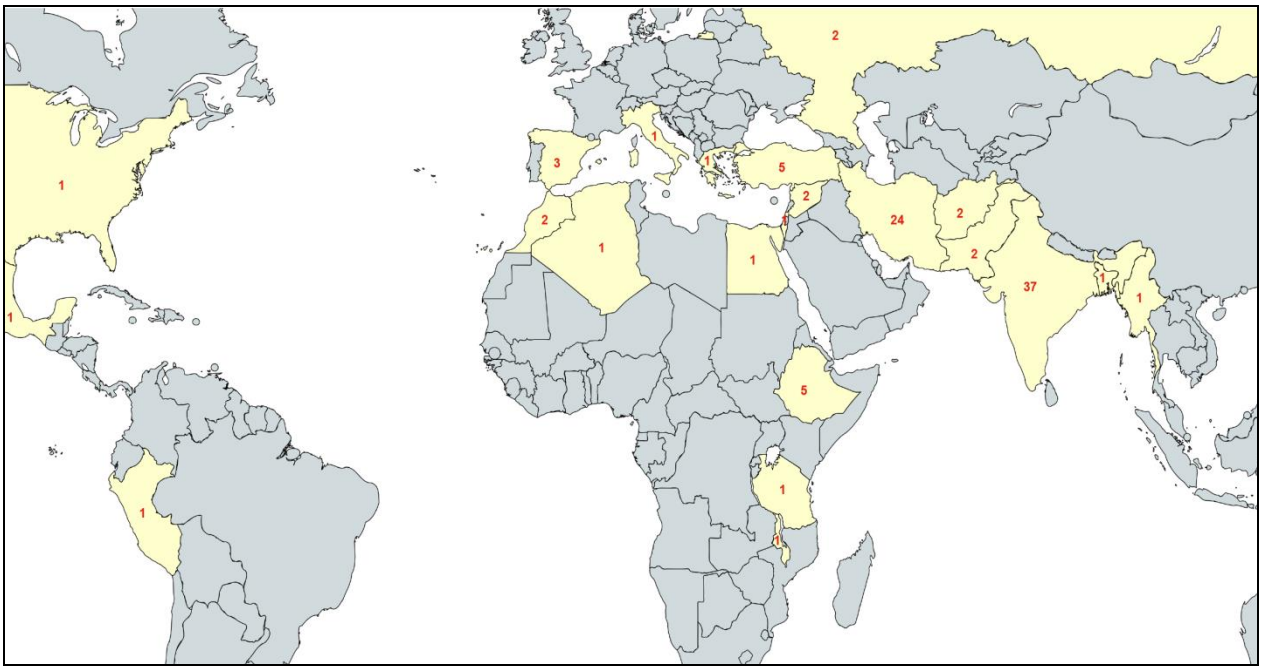


Figure 6.1 Geographical origin of the 96 chickpea lines used in the present study, indicating the number of accessions from each country.

To explore the possibility of extrapolating to field conditions (with more agronomical value) any polymorphism-phenotype association found under controlled conditions, we analysed the correlation between DTF recorded in the phytotron (Hobart) and field (Turretfield). A summary of the mean values obtained for the population in each of these environments as well as the correlation coefficient is shown in Table 6.3. For a full report of the flowering time recorded for each chickpea line, refer to appendix 6.2.

In general, the population was earlier to flower in Hobart than in Turretfield, even when it was grown under non-inductive SD (Table 6.3, Figure 6.2 B). This can be attributed to different photothermal conditions related to their difference in sowing dates; in the field the population was sown in autumn/winter, and thus experienced not only short photoperiod but also low temperatures during the first two months of growth (Turretfield research centre historical climate conditions are available at http://www.bom.gov.au/climate/averages/tables/cw_023343.shtml). By contrast, the population in Hobart was sown in January, coinciding with the maximum yearly temperatures, which are magnified by the glasshouse conditions.

In any case, a highly significant positive correlation was found between flowering times in both locations (Fig 6.2 A), which is especially strong when accessions were grown under short photoperiod (r = .675, p < 0.001). This correlation indicates that, despite a certain delay on the initiation of flowering as a consequence of environmental differences, the relative flowering behaviour of different accessions was preserved under field and glasshouse conditions.

Table 6.3 Mean and Standard deviation (S) values of flowering time obtained for the chickpea population grown in four different environments and their correlation, tested using Spearman's ranked correlation and indicating its statistical significance

	Mean	S	Spearman's rho	LD 2016	SD 2016	LD 2015	Turretfield
LD 2016	30.5	3.7	Correlation Coefficient	1.000	.533**	.773**	.367**
			Sig. (2-tailed)	.	.000	.000	.000
			N	96	96	87	87
SD 2016	61.7	26.2	Correlation Coefficient	.533**	1.000	.655**	.675**
			Sig. (2-tailed)	.000	.	.000	.000
			N	96	96	87	87
LD 2015	25.5	3.5	Correlation Coefficient	.773**	.655**	1.000	.543**
			Sig. (2-tailed)	.000	.000	.	.000
			N	87	87	87	87
Turretfield	98.9	8.7	Correlation Coefficient	.367**	.675**	.543**	1.000
			Sig. (2-tailed)	.000	.000	.000	.
			N	87	87	87	87

**. Correlation is significant at the 0.01 level (2-tailed).

As expected, significant and strong correlation was also found between DTF of plants grown in LD across seasons in Hobart (r = .773, p < 0.001). Also, positive and moderate to strong correlation was found between flowering time of the lines grown under different photoperiods (r = .655 and r = .533 in 2015 and 2016, respectively, with p < 0.001 in both cases). This consistency in the results across seasons and conditions indicates a high influence of the genetic component in the flowering of the lines.

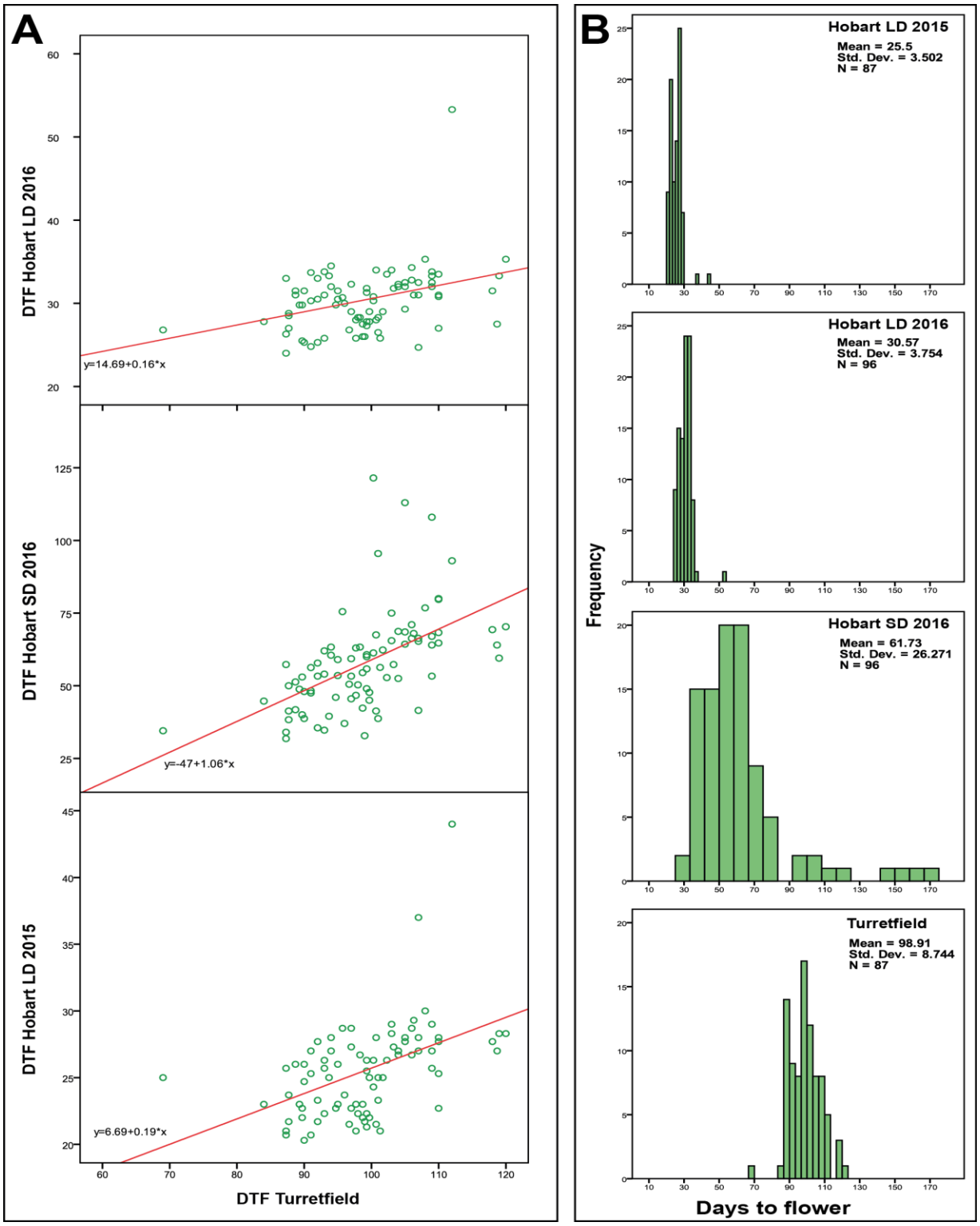


Figure 6.2 (**A**) Relation between flowering time of the chickpea population obtained in field (Turretfield, horizontal axis) and under 3 different environments in controlled environment facilities (Hobart, vertical axis). (**B**) Frequency distribution of the flowering time obtained for the same population in 4 environments; Turretfield, Hobart (2015) and Hobart (2016) under two photoperiods (LD =long days, SD = Short days).

6.3.2 FT cluster sequencing on a chickpea collection

The ~ 57 kb genomic region of chickpea chromosome 3 containing *FTa1*, *FTa2* and *FTc* was amplified in 6 independent PCR products (Fig 6.3 A), in the collection of 96 chickpea lines selected. Attempts to amplify various regions of *FTa2* were unsuccessful in 40 cultivated accessions (Fig 6.3 C, D), implying the presence of a large deletion in the region. Through the use of several different primer combinations, the approximate size of the deletion was defined as ~ 30 kb (Fig 6.3 E). As a result, a modified strategy with different primer combinations was used to amplify the cluster in those accessions carrying the deletion (Fig 6.3 B).

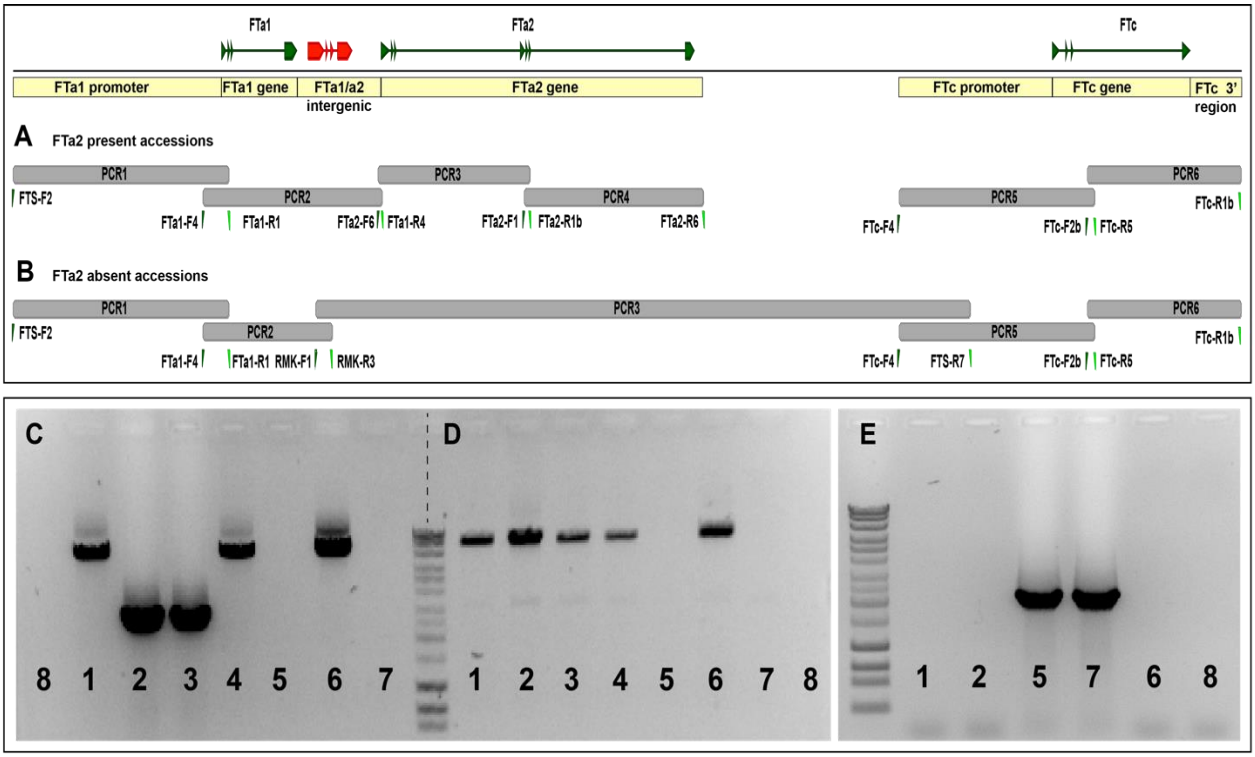


Figure 6.3 Schematic diagram of the region of chromosome 3 targeted for amplification and resequencing. The black line represents genomic DNA sequence, with *FTa1*, *FTa2* and *FTc* genes and relevant flanking regions indicated in yellow boxes. The *FTa1*, *FTa2* and *FTc* gene models are shown at top in green, and a non-coding RNA annotated in the *FTa1-a2* intergenic region (NCBI Gene ID 105851807) is indicated in red. The different PCR amplicons (grey boxes) and the position of the primers used (green triangles) are indicated below in accessions with an intact *FTa2* region (**A**) or those carrying the *FTa2* deletion (**B**). Forward and reverse primers are presented in dark and light green, respectively. Electrophoresis gel of the products obtained with primers FTa2-F6/FTa2-R1b (**C**), FTa2-F1/FTa2-R6 (**D**) and RMK-F1/FTS-R7 (**E**) in 7 chickpea accessions, as follows: 1-ICC4958; 2-PI489777; 3-Cr5-9; 4-CA2156; 5-WR315; 6-ILC3279; 7- ICCV2; 8-SDW. Expected size of the amplicon in (E) is 33.7 kb according to CDC Frontier genome, so the presence of a 2-3 kb band, indicates a probable deletion spanning ~ 30kb.

The PCR amplicons were purified and sequenced using Next-Generation Amplicon Sequencing as detailed in the Methods. The run report indicated good values for all quality metrics, including a density of flowcell clusters of 1295 ± 24 per mm^2, an 86.49 ± 0.83 cluster PF (percentage of clusters that passed Illumina's so-called "chastity filter" and can provide a clear signal for the sequencing process) and very low value of Phasing/Prephasing (0.0435/0.0415). This indicates the percentage of true signal being lost in each cycle due to desynchronization of individual molecules in a cluster with others by either phasing (they fall behind) or pre-phasing (jumping ahead). The sequencing process yielded a total of 6.6 Gbp distributed in 48311712 reads, with an average of 503247/reads per accession (the proportion of reads assigned to each genotype is indicated in appendix 6.1). The majority of these reads (90.91 %) were high-quality reads, with a score equal or higher than Q30 in the Phred quality score.

To obtain final sequences of the *FTa1-FTa2-FTc* genes in each of the 96 genotypes, the data was processed using two different approaches. These consisted of 1) mapping the reads against the reference genome sequence from cultivar CDC Frontier available on NCBI, and 2) a *de novo* assembly of the reads for each accession. Mapping against a reference genome can easily identify SNPs and small insertion or deletions (indels) between any accession and the reference sequence, but this strategy does not detect large indel events. *De novo* assembly was therefore the method of choice to detect repetitive sequence and large indels. Although this dual approach was more time consuming, the combination of methods ultimately resulted in a more accurate retrieval of all polymorphisms present.

6.3.3 Interspecific sequence differences in the FT cluster

The *FT* cluster sequences from both *C. reticulatum* accessions used in this study (PI489777 and Cr5-9) were almost identical. A single polymorphism (G/A) was found between both sequences in the third intron of *FTa2* among the 50736 bp sequenced. Therefore, only the sequence from accession PI489777 was used. This sequence was initially compared with that from the *desi* cultivar ICC4958, as these two lines are parents of the CRIL2 reference cross (Sharma et al. 2013) in which the major QTL in the *FT* cluster region was defined (Chapter 4).

The alignment between the sequences from PI489777 and ICC4958 (57048 bp long) revealed a percentage of homology of 87.0% and a high degree of polymorphism, with numerous variable

sites observed across the alignment and a total of 365 SNPs and 73 indels (of variable length) identified (Fig 6.4). A fully detailed list with the position and sequence of all the polymorphisms is available in Electronic Supplementary Material 2.

To further investigate this variability, the *FT* cluster was divided into 7 different regions, as follows (represented by yellow boxes in Fig 6.3); the *FTa1* promoter (beginning 11.5 kb upstream from the start of *FTa1* mRNA), the *FTa1* gene (corresponding to the genomic sequence of *FTa1* including exons and introns), the *FTa1-a2* intergenic region, *FTa2* gene (exons and introns), *FTc* promoter (from ≈8 kb upstream to the start of *FTc* mRNA), *FTc* gene (exons and introns) and *FTc* 3' region (4.6kb after the end of *FTc* mRNA). The nucleotide (π) and indel (π_i) diversity as well as the number of SNPs and indel events were for calculated for each of these regions and across the cluster as a whole (Table 6.4). The *FTa1* promoter was found to be the most variable by far, with a nucleotide diversity of 0.02176 and 70.6 % of the total SNPs and 41 % of the indels. This contrasts with the other promoter region analysed (*FTc*), which in comparison showed much lower values for both nucleotide (0.0035) and 7 % of the total SNPs and 10 % of the indels. The region corresponding to the *FTa2* gene was the most conserved, displaying the lowest diversity values of the seven regions comprising the cluster.

Despite the elevated number of polymorphic sites, the coding regions were highly conserved; only two SNPs were found in the coding sequence of the three *FT* genes. The first is a silent substitution (T/C) in the first exon of *FTa1*, located 60 bases after the ATG codon. The second, a (G269T) transversion, affects the last nucleotide of the second exon and introduces the amino acid change Trp90Leu.

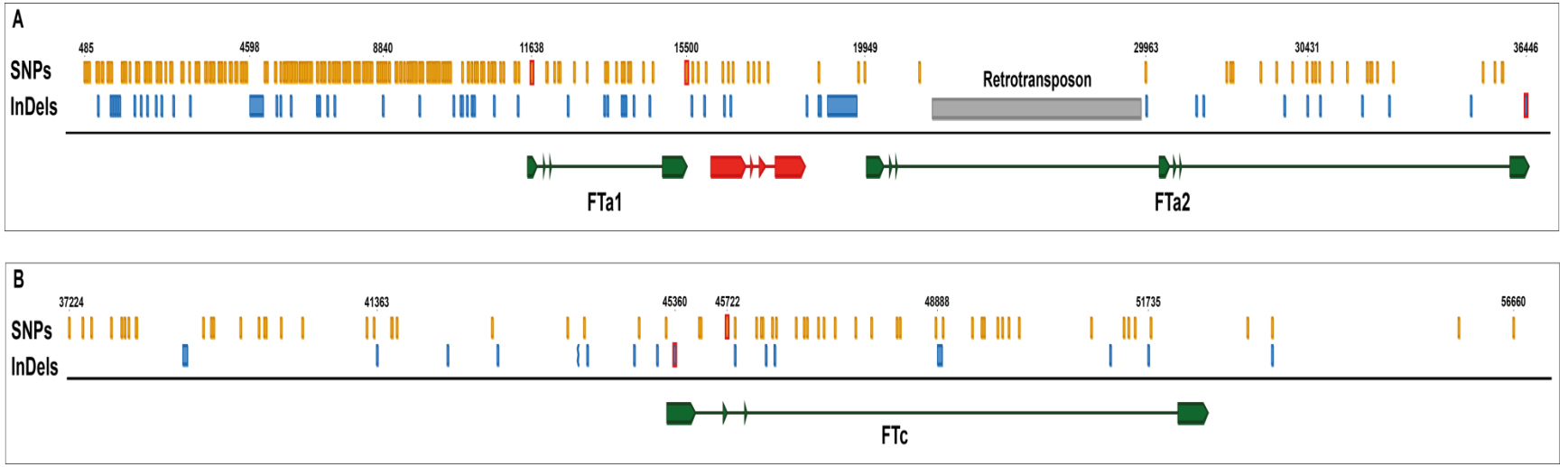


Figure 6.4 Diagram of the *FTa1-FTa2* cluster (**A**) and the *FTc* gene (**B**) showing the position of all SNPs (orange lines) and Insertion-Deletions (indels, blue lines) found between sequences obtained from wild accession PI489777 and *desi* cultivar ICC4958. SNPs and indels in coding or UTR regions in any of the three genes are highlighted in red. mRNA of the different genes is represented by green boxes (exons) and lines (introns) while the ncRNA is indicated in red following the same scheme. A retrotransposon found in the third *FTa2* intron in ICC4958 is depicted as a grey rectangle.

Table 6.4 Summary of nucleotide diversity (π) and indel diversity (π_i) parameters as well as the number of SNPs and indel events found among the different regions within the *FT* cluster of *C. arietinum* and *C. reticulatum*. Since only 2 sequences were used, θ and π values are similar and therefore only π is shown.

	FTa1			FTa1/a2 Intergenic region	FTa2	FTc			Total
	Promoter	Gene[a]	Total		Gene[a]	Promoter	Gene[a]	3' region	Total	
Size (bp)	11526	3991	15517	4466	17039	8075	7348	4603	20026	57048
Nucleotide diversity (π)	0.02176	0.00603	0.01753	0.00352	0.00212	0.0035	0.00515	0.0013	0.00359	0.00736
Number of SNPs	234	24	258	13	25	28	35	6	69	365
Number of indels	30	8	38	8	11	8	7	1	16	73
Indel diversity	0.00269	0.002	0.00251	0.00202	0.00065	0.00099	0.00109	0.00043	0.0009	0.00135

a) Genomic region comprising transcribed exons and introns

Polymorphisms were also found in the untranslated regions (UTR) of the different *FT* genes. A SNP (T/C) was found in the 3' UTR of *FTa1*, very close to the end of the mRNA. Also in the 3'UTR of *FTa2* a 1-base indel was found in a string of 10 T (the cultivated allele possesses an extra T). The 5' UTR of the *FTc* mRNA contains a microsatellite (AT repetition) variable between alleles; the wild accessions possess 11 repetitions of this motif whereas cultivated accession has 13 repetitions. Since there is no way to predict the functional implications of these changes, further research would be needed to test this.

FTa2 retrotransposon insertion.

According to the annotation of *CaFT2* gene in NCBI (Gene ID 101496618), this gene possesses a distinct and unusual genomic structure: Whereas *FT* genes in general have a conserved genomic structure with 3 introns in the coding region, *FTa2* has an extended 5'UTR containing three additional introns (see Chapter 3, section 3.3.1). The most significant difference observed between the *FT* clusters of accessions PI489777 and ICC4958 was a large insertion of 5219 bp located just 508 bp upstream of the *FTa2* coding region in ICC4958 (Fig. 6.4), and situated within the third of these *FTa2* additional introns. This insertion contained a large open reading frame (ORF) of 4521 bp encoding a protein with several domains characteristic of a retroelement (retrotansposon or retrovirus); a reverse transcriptase, an integrase (that mediates integration of a DNA copy of the viral genome into the host chromosome), a *gag* gene (encodes structural proteins that form the virus-like particle inside which reverse transcription takes place) and a ribonuclease H (responsible of the original RNA template degradation, generation of polypurine tract and final removal of RNA primers from newly synthesized DNA strands). The presence of Long Terminal Repeats (LTR) and the phylogenetic relationships of the different domains identified this element as a LTR-retrotransposon of the *Ty1/Copia* type (Finnegan 2012; Ustyantsev et al. 2015).

In addition to this element (from now referred as Retrotransposon 1, or RT1), the ORF analysis identified two more retrotransposons within the *FT* cluster, a second in the *FTa2-FTc* intergenic region (RT2, 5265 bp) and a third (RT3, 5319 bp) in the third intron of *FTc* (Fig 6.5). Both RT2 and RT3 possess LTRs flanking their coding regions and have similar domains to those present in RT1, so it is assumed that they all belong to the same type of retrotransposon (for further

information about the position, BLAST E-values and a brief description of the different domains present in each retrotransposon refer to appendix 6.3.

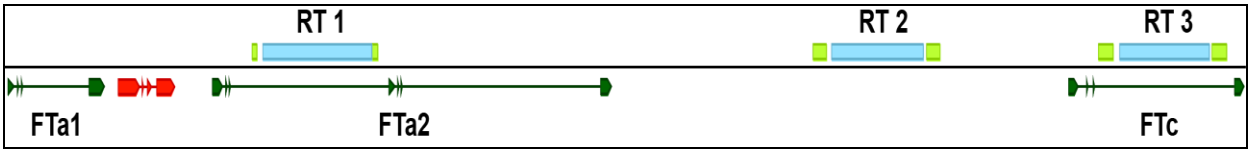


Figure 6.5 Schematic representation of the *FTa1-FTa2-FTc* cluster showing the position of the three retrotransposons (RT) found within it. mRNA from *FT* genes are in green, ncRNA in red, with arrows representing exons and lines introns. The coding regions of the RT are indicated by blue boxes and LTRs by light green boxes.

Both RT2 and RT3 insertions are present in both PI489777 and ICC4958 indicating they took place prior to the domestication process and therefore cannot be the cause of any differential behaviour observed between the two species. RT1, on the other hand, is present only in the cultivated lines, which suggests a domestication-related event.

Insertions of transposable elements can influence gene regulation in multiple ways (Galindo-González et al. 2017; Cui and Cao 2015), so in the next section, I will investigate the possibility of mRNA alteration.

FTa2 alternative splicing

To confirm that the unusual transcript of *FTa2* was indeed expressed, and to test whether the RT1 insertion might somehow modify its splicing pattern, different primers combination targeting *FTa2* were tested on the cDNA of both wild and cultivated accession (Fig 6.6).

As expected, a single band was obtained for several different primer combinations within the coding region when cDNA was used as template (Fig 6.6, B and C). However, several bands were observed when the forward primer was placed in any of the additional exons within the 5'UTR (Fig 6.6 D and E). A pattern consisting of two equivalently strong bands differing in size by ≈200 bp plus several other larger, fainter bands was found in leaves of both cultivated (ICC4958) and wild accessions (PI489777), and confirmed in two more accessions, one of each species (cultivated ICCL81001 and wild Cr5-9, data not shown). The pattern was independent of age or photoperiod with the conditions used in this study (Figure 6.6 D and E). These results

confirm the integrity of the annotated *CaFTa2* structure and also show that the splicing pattern of this gene is not affected by the presence of the retrotransposon insertion RT1.

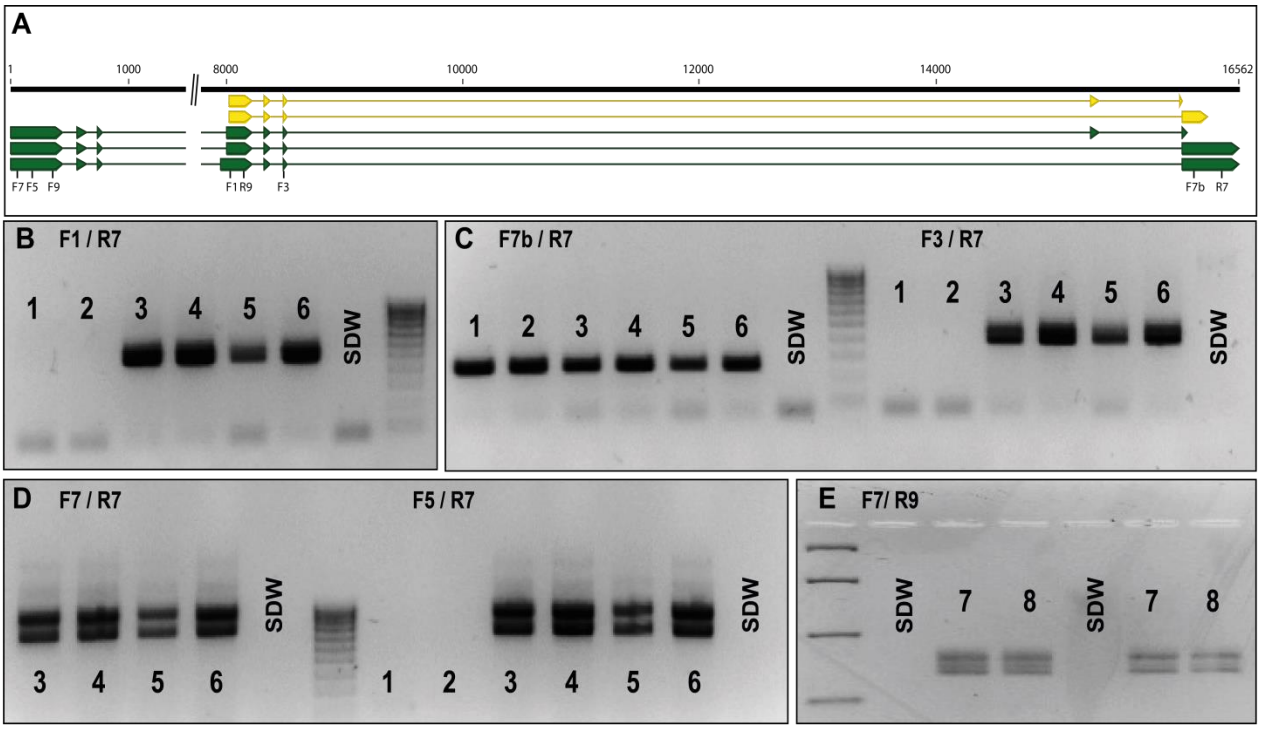


Figure 6.6 (**A**) Representation of the *FTa2* genomic sequence and the position of primers used. mRNA and coding region are indicated by green and yellow boxes, respectively. Length of the intron 3 (containing the retrotransposon, which is not shown for this reason) has been reduced for graphical representation. (**B** to **E**) Visualization of the PCR products obtained with different primer pairs. Hyperladder IV (B to D) or Easy ladder I (E) were used as size indicator (Bioline). Lanes 1 and 2- Genomic DNA from accession ICC4958 (1) and PI489777 (2); Lanes 3 to 7 - cDNA from ICC4958 leaves harvested when plants were four (3, 5), six (4, 6) or eight (7) weeks old. Lane 8- cDNA from PI489777 leaves harvested from twelve weeks old plants. Plants were grown in either long days (16 h photoperiod, lanes 3, 4, 7 and 8) or short days (8 h photoperiod, lanes 5 and 6).

To further investigate the identity of the multiple bands observed in both species, the PCR products obtained using the primers F7/R7 were cloned only from the accession ICC4958 and sequenced. Alternative splicing of this gene was already predicted by sequence data in NCBI, with 3 different mRNA isoforms described (Fig 6.7). The analysis of sequences from ICC4958 revealed the presence of two of these isoforms, and 8 other distinct transcripts. The most significant changes compared with the canonical transcript were loss of the fourth (first coding) exon, and the inclusion of new exons between exons 6 and 7 (Figure 6.7). According to NCBI, the last exons present a long (484 bp) and a short variant (55 bp). We detected 8 novel transcripts

using a primer combination that amplify only those spliced forms with the long variant. This means that the number of *FTa2* isoforms is higher than predicted, as other variants could be obtained with a different primer combination.

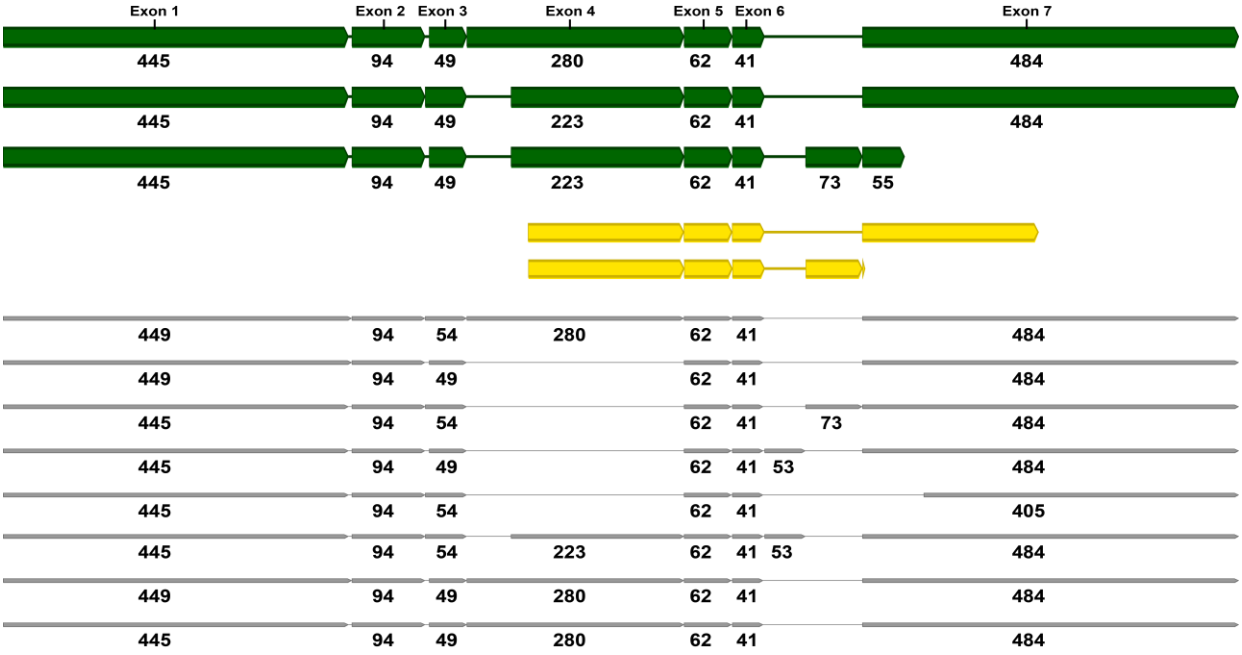


Figure 6.7 Alternative splicing of chickpea *FTa2*. The 3 different mRNAs and coding regions annotated in NCBI are indicated in green and yellow, respectively. Eleven different mRNAs were sequenced in this study, which are indicated in grey. Numbers over the exons correspond to size in bp.

6.3.4 Diversity among cultivated accessions

Diversity within the collection of domesticated chickpea accessions was then characterized. Only one major problem was encountered, involving the amplification and sequencing of fragment PCR5 (promoter and 5' end of *FTc* gene; Fig 6.3), which was successful in only 22 out of 94 accessions. After analysis of the PCR5 sequence from ICC4958, it became apparent that the primers designed to amplify PCR5 were unfortunately located within RT2 (forward primer) and RT3 (reverse primer). In many accessions this primer pair unfortunately amplified an alternative non-specific product of the same approximate size as the one expected. However, the adjacent fragment (PCR6) covering the last exon of the *FTc* gene and 4.5 kb of the 3' downstream flanking sequence was successfully sequenced in the whole population, thus providing at least partial *FTc* sequence for analysis (Fig 6.9).

In contrast to the high level of polymorphism seen in the wild/domesticated comparison, a much lower level was found among the 94 domesticated accessions. Table 6.5 shows that all nucleotide diversity parameters displayed smaller values than those obtained when comparing alleles from wild and cultivated lines (π values of 0.00736/0.00034 and π_i values of 0.00135/0.00059 for inter/intraspecific alignments, respectively).

Despite the lower genetic diversity, 66 SNPs and 26 indels can be found across the FT cluster in different chickpea accessions (Table 6.5). Further information about the position, size and sequence of these changes can be found in Table 6.6.

Table 6.5 Summary of nucleotide (π and θ per site) and indel diversity (π_i) parameters as well as the number of SNPs and indel events found in the *FT* cluster among a collection of *C. arietinum* accessions

	FTa1			FTa1/FTa2 Intergenic	FTa2	FTc	Total
	Promoter	Gene[a]	Total		Gene[a]	Total	
Size	11002	3994	14996	4464	17113	10002	46575
Nucleotide diversity (π)	0.00012	0.00002	0.00009	0.00240	0.00014	0.0001	0.00034
θ per site	0.00045	0.0001	0.00035	0.00427	0.0003	0.00039	0.00072
Number of SNPs	22	2	24	1	21	20	66
Number of indels	4	2	6	7	10	3	26
Indel diversity (π_i)	0.00027	0.00001	0.00019	0.00062	0.00026	0.00002	0.00059
Haplotype number							
Gaps not considered	14	3	16	1	16	18	28
Gaps considered	83	7	85	22	54	21	93
Haplotype diversity							
Gaps not considered	0.4496	0.083	0.4976	-	0.6961	0.5895	0.00059
Gaps considered	0.9974	0.1631	0.9981	0.8582	1	0.6421	0.9998

Genomic region comprising transcribed exons and introns

Figure 6.8 and Figure 6.9 show the position of the polymorphisms found in the *FT* cluster among the 94 cultivated lines. With the exception of three SNP observed in the case of *FTc*, the coding sequence of the three genes was highly conserved. *FTa1* displayed the lowest level of polymorphism among the three genes; not a single SNP was found in the transcribed region, and only 2 SNPs and 2 indels were found within the introns, contrasting with the much higher number of polymorphic sites observed in *FTa2* and *FTc* introns. Moreover, there is a region of 1300 bp before the *FTa1* mRNA start point that is also highly conserved, with no polymorphism was detected across the accessions analyzed. The overall high degree of conservation across the

FTa1 gene is consistent with an important role for this gene. The UTR regions show also high overall levels of conservation, although in the case of *FTa2* some variation was found in two accessions; ICC5878 has a SNP in the first *FTa2* exon, in the extended 5' UTR, while ICC15435 carries a 48 bp insertion in the 3' UTR. Unfortunately, whether these variants have any effect on gene function is again difficult to predict. As expected, most of the variations were located within non-coding regions (Fig 6.8, Fig 6.9 B). The most obvious of these were an insertion of ~750 bp in the intergenic region between *FTa1* and *FTa2*, very close to the transcription start site of the *FTa2* gene (Fig 6.8 C), and a large deletion of ~ 30 kb starting in the *FTa1*/*a2* intergenic region and covering the whole genomic sequence of *FTa2* gene and part of the *FTa2*/*FTc* intergenic region (Fig 6.8 B). I initially thought that this was a single deletion common to all accession lacking *FTa2*, but once the *FT* cluster was sequenced in all lines it became clear that there were two different forms, which can be distinguished as Type 1 and Type 2 (Figure 6.8) Although similar in size, both start (Fig 6.8 C) and end points (Fig 6.8 B) of the two deletions are different, (Table 6.6), indicating that they are the result of independent events.

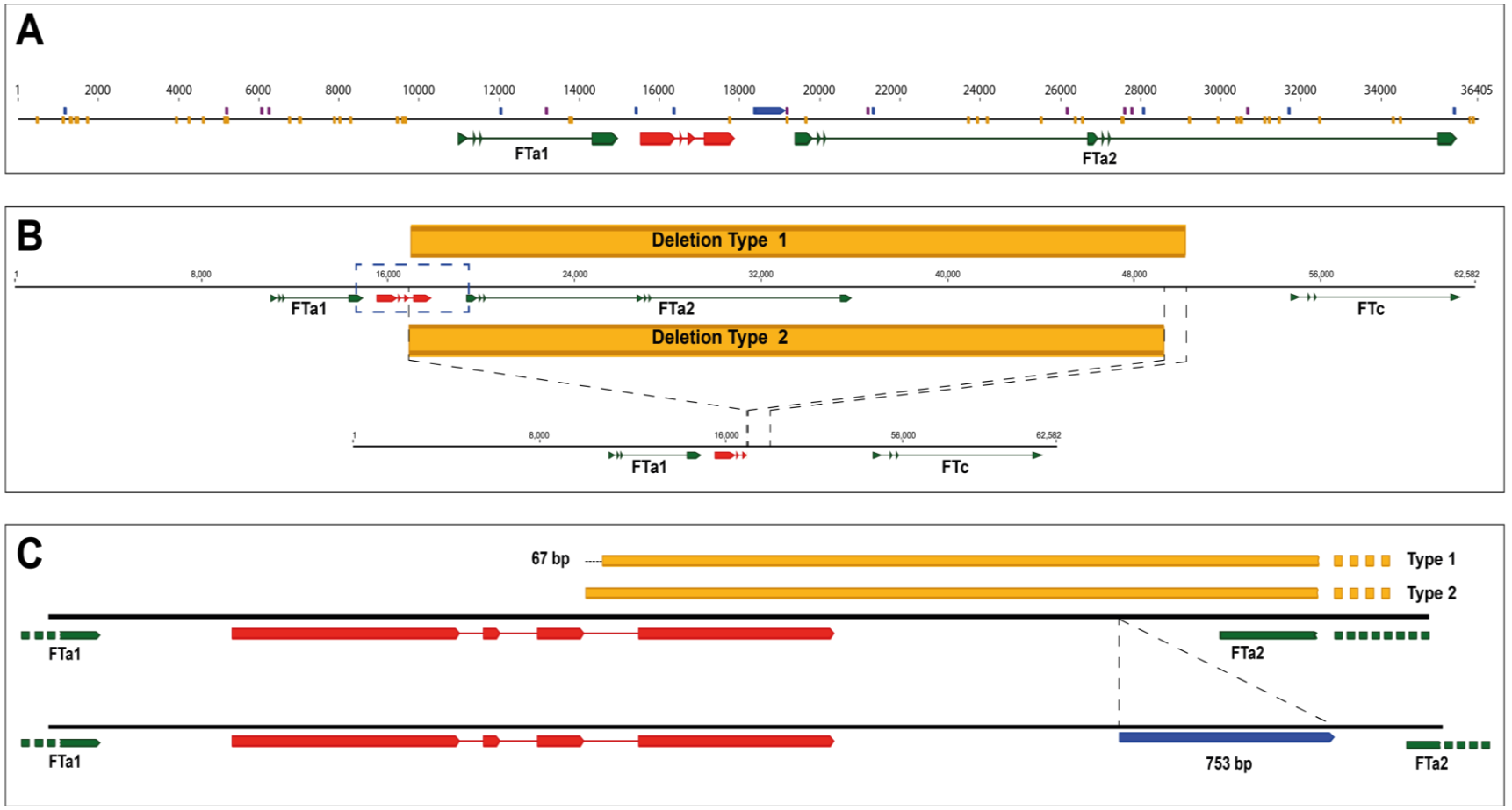


Figure 6.8 (**A**) Polymorphisms found in the *FTa1-FTa2* cluster among 94 chickpea cultivated accessions. SNPs are indicated in orange, indel events in blue and microsatellites in purple. (**B**) Schematic diagram showing the extent of *type 1* (33238 bp) and *type 2* (32385 bp) deletions and the resulting *FTa1-FTc* cluster. The region squared with dashed blue lines correspond to the *FTa1-a2* intergenic region, zoomed in (C). (**C**) *FTa1-a2* intergenic region showing the 67 bp difference in the start point of deletions *type 1* and *2*, as well as the location of the ~750 bp insertion found in some accessions (blue box). In all cases, black line represents the genomic DNA, and the mRNA of the different genes are depicted in green under the genomic DNA, with boxes representing exons and line introns. The ncRNA annotated in the *FTa1-a2* intergenic region is shown in red, following the same pattern. Numbers indicate distance in base-pairs.

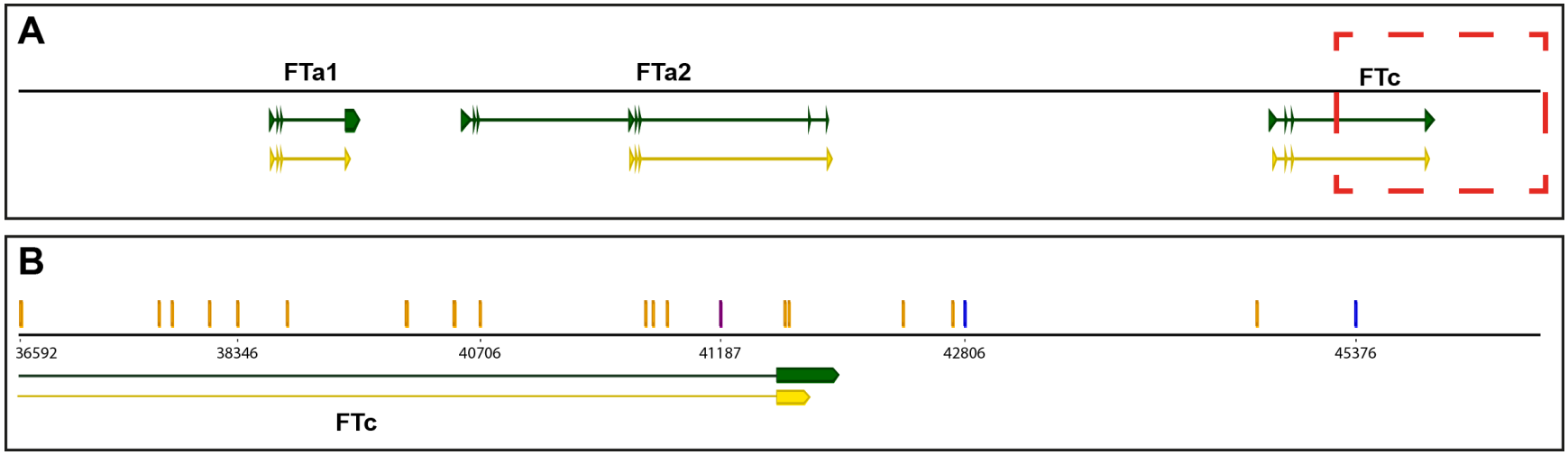


Figure 6.9 (**A**) Schematic diagram of the *FTa1-a2-c* cluster. The red (dashed) square indicates the region of *FTc* that was successfully sequenced in 94 cultivated chickpea accessions (shown amplified in figure 6.9-B)**.** (**B**) Polymorphisms found in the portion of *FTc* gene and 4597 bp of the 3' region sequenced in 94 chickpea cultivated accessions. SNPs are indicated in orange, indel events in blue and microsatellites in purple. mRNA of the different genes are depicted in green, and the coding region in yellow, with boxes representing exons and line introns. Numbers indicates distance in base-pairs.

Table 6.6 Position, length and sequence of the different single nucleotide polymorphism (SNP), insertion and deletion events (indel) and microsatellites (MS) found in the alignment of the *FTa1-a2-c* sequences from a collection of 94 *C. arietinum* accessions. When more than 2 variants found, the number of alleles is also indicated. The position indicated is relative to the alignment, so these numbers can vary when the sequences are individually examined.

Position	Feature	length (bp)	Polymorphism	Alleles
479	**SNP1**	-	C/A	-
1130	**SNP2**	-	T/C	-
1174	**Indel1**	1	T/-	-
1320	**SNP3**	-	T/C	-
1339	**SNP4**	-	A/G	-
1459	**SNP5**	-	G/A	-

Position	Feature	length (bp)	Polymorphism	Alleles
1489	**SNP6**	-	C/T	-
1741	**SNP7**	-	C/G	-
3958	**SNP8**	-	C/T	-
4264	**SNP9**	-	C/A	-
4624	**SNP10**	-	C/T	-
5160	**SNP11**	-	G/A	-

Table 6.6 Continued

Position	Feature	length (bp)	Polymorphism	Alleles	Position	Feature	length (bp)	Polymorphism	Alleles
5180	**SNP12**	-	C/T	-	17008-36572	**Del Type 1**	19565	-	-
5195-5211	**MS1**	17	G repeat	6	19737	**SNP26**	-	C/T	-
5233	**SNP13**	-	C/A	-	21274-21278	**MS6**	14	TA repeat	3
6074-6120	**MS2**	47	TA repeat	14	21409-21444	**Indel5**	36	b	-
6275-6310	**MS3**	36	TA repeat	7	23797	**SNP27**	-	G/C	-
6822	**SNP14**	-	A/T	-	24017	**SNP28**	-	C/A	-
7062	**SNP15**	-	T/C	-	24275	**SNP29**	-	G/A	-
7080	**SNP16**	-	A/G	-	25608	**SNP30**	-	G/A	-
7930	**SNP17**	-	G/C	-	26250-26267	**MS7**	18	TA repeat	3
8068	**SNP18**	-	G/A	-	26466	**SNP31**	-	C/T	-
8349	**SNP19**	-	C/T	-	26641	**SNP32**	-	A/G	-
9510	**SNP20**	-	C/A	-	27648	**SNP33**	-	C/T	-
9645	**SNP21**	-	A/G	-	27654	**SNP34**	-	C/T	-
9699	**SNP22**	-	C/T	-	27657-27752	**MS8**	96	TAA repeat	13
12086-12094	**Indel2**	9	TGGGCAAAG/-	-	27893-27906	**MS9**	14	T repeat	2
13224-13236	**MS4**	13	A repeat	3	28194	**Indel6**	1	T/-	-
13822	**SNP23**	-	C/T	-	29329	**SNP35**	-	T/C	-
13870	**SNP24**	-	A/G	-	30045	**SNP36**	-	C/T	-
15481	**Indel3**	1	A/-	-	30520	**SNP37**	-	T/C	-
16425	**Indel4**	1	C/-	-	30623	**SNP38**	-	G/A	-
17811	**SNP25**	-	A/T	-	30630	**SNP39**	-	G/A	-
18643-19215	**753 bp Indel**	753	a	3	30738-30839	**MS10**	102	TAA repeat	14
19216-19292	**MS5**	77	TTA repeat	10	31248	**SNP40**	-	T/A	-
16941-36572	**Del Type 2**	19632	-	-	31369	**SNP41**	-	T/A	-

Table 6.6 Continued

Position	Feature	length (bp)	Polymorphism	Alleles
31596	**SNP42**	-	G/A	-
31858-31861	**Indel7**	4	TCGA	-
32618	**SNP43**	-	C/A	-
34456	**SNP44**	-	T/G	-
34640	**SNP45**	-	C/G	-
35965-36012	**Indel8**	48	c	-
36410	**Indel9**	1	A/-	-
36500	**SNP46**	-	A/T	-
36592	**SNP47**	-	T/C	-
36600	**SNP48**	-	A/G	-
37503	**SNP49**	-	C/A	-
37590	**SNP50**	-	C/A	-
37837	**SNP51**	-	T/A	-
38023	**SNP52**	-	T/G	-
38346	**SNP53**	-	C/T	-
39129	**SNP54**	-	G/A	-
39134	**SNP55**	-	G/A	-
39446	**SNP56**	-	G/A	-
39451	**SNP57**	-	C/A	-
39616	**SNP58**	-	G/A	-
40706	**SNP59**	-	C/T	-
40753	**SNP60**	-	C/T	-
40845	**SNP61**	-	G/A	-
41187-41210	**MS11**	24	TAA repeat	3
41641	**SNP62**	-	A/T	-
41650	**SNP63**	-	G/T	-
42400	**SNP64**	-	A/G	-
42729	**SNP65**	-	T/A	-
42806-42809	**Indel10**	4	AATA	-
44729	**SNP66**	-	C/T	-
45376	**Indel11**	-	G/-	-

a) The complete sequence of the insertion can be found in Electronic Supplementary Material 3
b) TCATTGCTGCGCCACCTTTAAGCATTGTTGCAGCCT
c) TTTTTAAGTTTGCACTTTAGATTAGAATAAGTACGTCCCCTGTTTTTT

6.3.5 Potential functional significance of major sequence variants

Three groups of polymorphisms were further examined for their potential functional significance, based on their location and potential to affect gene function or expression: 1) the *FTa2* deletions; 2) the 753 bp insertion in the *FTa1*/*a2* intergenic region, and 3) the three SNPs in the *FTc* coding region.

Effect of FTa2 deletions.

In the present study, 40 out of 94 of the *C. arietinum* accessions analysed (i.e. 42.5%) carried a total deletion of *FTa2* gene (either *type 1* or *type 2*). To test the distribution of this deletion within cultivated germplasm, a dominant marker was developed and used to screen an additional collection of 109 domesticated lines (performed by collaborators at University of Saskatoon, Canada). A list of all the accessions as well as flowering data in two different environments can be found in appendix 6.4. Forty-one of these accessions (38%) lacked *FTa2*, a proportion very similar to that seen in the selection from the ICRISAT reference set analyzed in Hobart, suggesting that around 40% of all chickpea cultivated germplasm lack the *FTa2* gene.

The *type 2* deletion was more common than *type 1*, present in 29 and 11 accessions, respectively. The proportion of *type 1* was similar between *desi* and *kabuli* accessions (11.1% and 13.6%, respectively), whereas the incidence of *type 2* seems to be higher between *desi* (37.5 % of the total accessions) than in *kabuli* lines (9.1%). The absence of *FTa2* seems to be more common within the South-west Asian germplasm, as the majority of the accessions carrying one or the other deletion were found in India or Iran (Fig 6.10 A).

To test the possible effect of the deletion on flowering, the mean flowering time (DTF, Days from emergence to first open flower) in the three groups (wild-type/*FTa2*$^{+}$, *type 1* or *type 2*) was compared in three different environments (Fig 6.10 B, C). Accessions carrying either deletion were always slightly later to flower than *FTa2*$^{+}$ lines when grown in LD (3.2 and 3.6 days for type 1, 1.9 and 2.4 days for type 2, in 2015 and 2016 respectively). This effect was also apparent when accessions were sorted according to their flowering time; lines with the deletions tend to group towards the later part of the distribution in both seasons, particularly evident for the type 1 lines (Fig 6.10 C). Under SD, *Type1* accessions are also later flowering whereas the type 2 deletion in contrast seems to be associated with earlier flowering under these conditions.

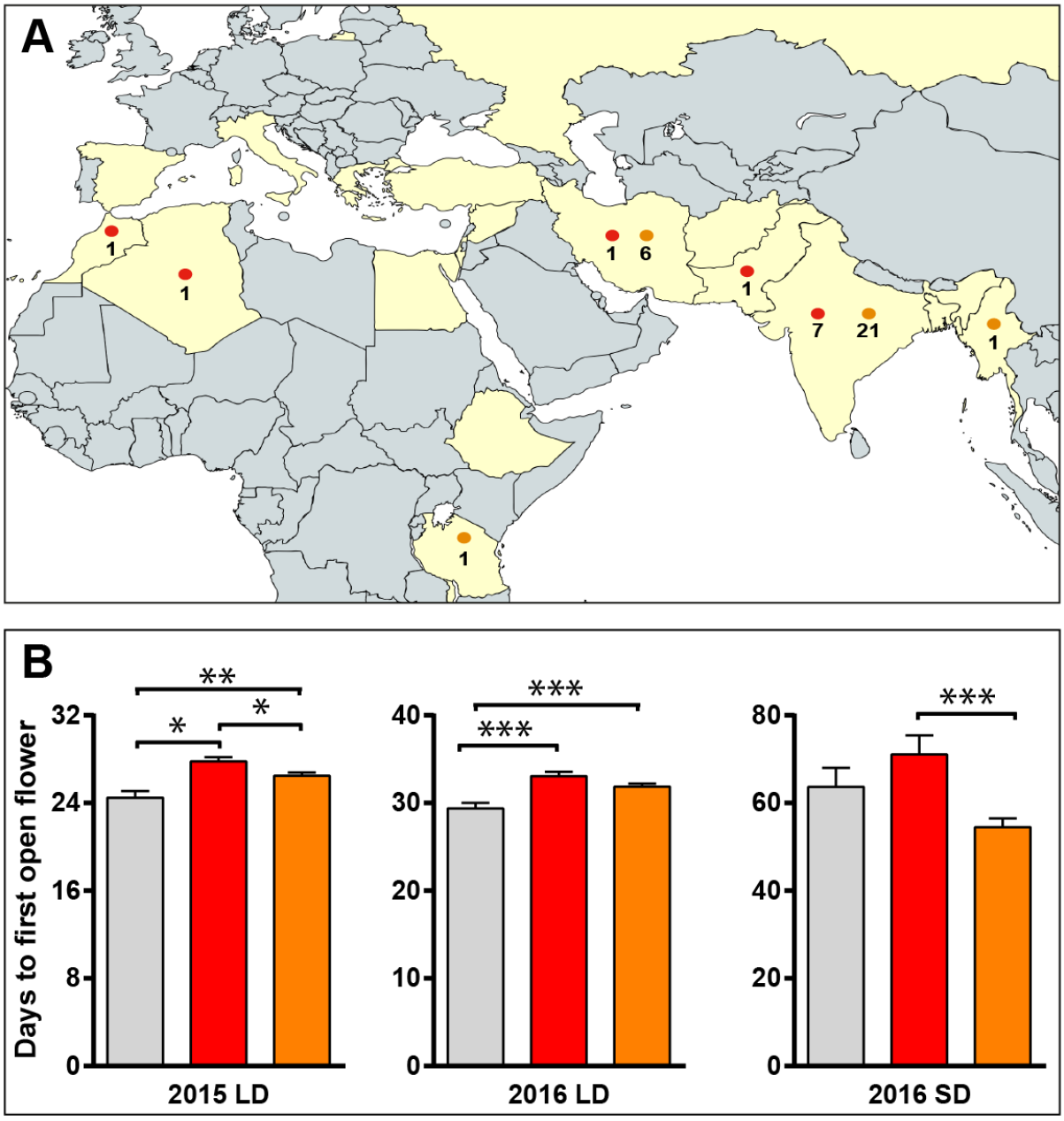
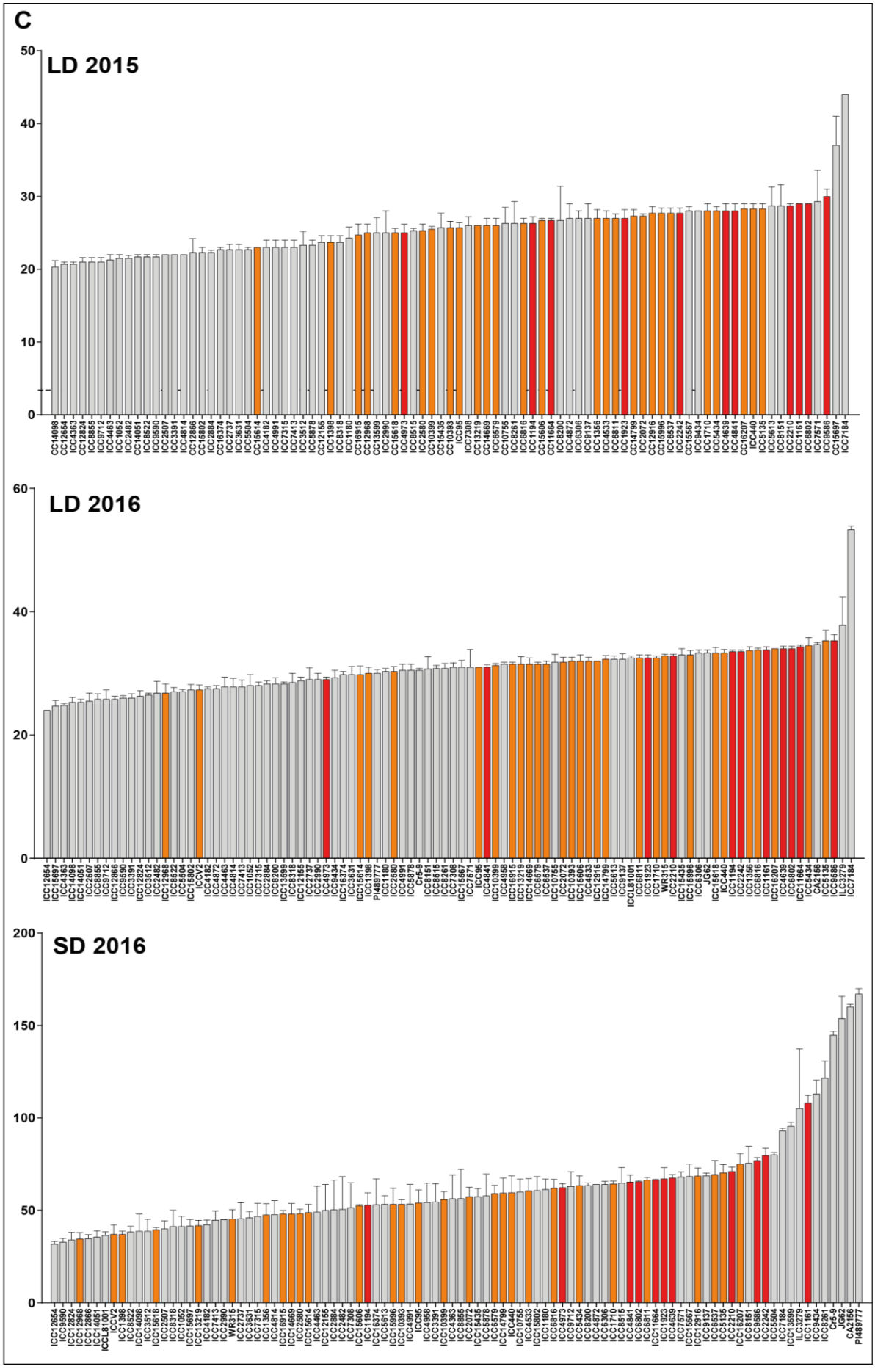


Figure 6.10 (**A**) Geographic distribution of the 40 chickpea lines with *FTa2* deletion *type 1* (red circles) or *type 2* (orange circles). (**B**) Days from emergence to first open flower of the cultivated lines grown in 3 different environments and pooled according to their *FTa2* genotype: *FTa2* present (grey bars), *FTa2* deletion *type 1* (red) or *type 2* (orange). Asterisk denotes significance level (* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$) (**C**) (on next page) Mean DTF recorded in Hobart for each of the 94 cultivated lines grown in three different environments. Accessions were sorted from earliest to latest. Colour of the columns corresponds to *FTa2* genotype, as described in (B).


C
LD 2015
LD 2016
SD 2016

Effect of the 750 bp insertion upstream of FTa2

Among the 54 cultivated accessions in which the *FTa2* gene was present, 10 lines were found to have an insertion of around 750 bp, which itself showed minor sequence/ length variation across the 10 lines. These ten lines flowered significantly earlier (50.5 ± 2.4) compared to the 44 lines without the insertion (62.4 ± 4.4) when grown under SD ($t(51) = 2.334$, $p = 0.24$), but there was no difference between the two groups under LD.

SNPs in FTc coding region

Three different SNPs were found in the *FTc* coding region; one in exon 2 and two in exon 4. The first (G269T) affected the last nucleotide in exon 2 (mentioned above in section 6.3.3). This is a non-synonymous substitution that specifies a Trp90Leu change in the FTc protein. Interestingly, this residue corresponds to Trp88 in Arabidopsis FT (AtFT) which is located very close to a key residue Tyr85 determining FT activity (Hanzawa et al. 2005). This SNP is not included in Table 6.6, since as previously stated, the sequencing of the first half of *FTc* was successful in only in 22 accessions of the population (two wild lines and 20 cultivated). Among these 22 accessions, two wild and 7 cultivated lines carried the G allele and the other 13 cultivated lines a T.

Two more non-synonymous mutations were identified in the last exon of *FTc*: line ICC4363 carries a SNP (G395T, SNP63 in Table 6.6) that introduces the amino acid substitution Arg132Met (Arg130Met in AtFT protein), while the accession ICC9434 carries a transversion (A386T, SNP62 in Table 6.6) directing a Glu129Leu substitution (Glu127Leu in AtFT). These two amino acid changes are adjacent to or included in the so called "Segment B", a region of the FT protein (positions 128 to 141) that has been identified as critical for the function of this gene (Ahn et al. 2006). Segment B residues form an external loop, as determined by analysis of crystal structures. This is, therefore, a distinct part of the protein to that previously mentioned for Tyr85, but both are needed for a correct formation of the ligand-binding site (Hanzawa et al. 2005; Ahn et al. 2006).

To investigate the conservation of these three amino acids, we aligned the proteins from the wild *CrFTc* gene (deduced from the sequence obtained from accession PI489777), CaFTc (from accession ICC4958), AtFT (used as reference and evaluate the conserved regions) and 57 other PEBP proteins from several different legumes. Figure 6.11 shows that Gln127 was highly conserved across all PEBP proteins. Analysis of segment B in previous studies show

that Leu128 (in the edge of the segment) is variable, with different amino acids occupying this position in different proteins. Therefore, if changes at this site are tolerated, it is possible that this might be extensible to neighbouring places (Wickland and Hanzawa 2015). Arg130 is conserved among FT/TFL1 and BFT proteins in the legumes analysed, but is replaced by a proline residue in the MFT proteins. Previous reports indicated that the substitution Arg130Met could cause an early flowering phenotype in Arabidopsis (Ho and Weigel 2014). Interestingly, the accession where this mutation was detected (ICC4363) was the third earliest accession when grown under LD in both 2015 (with a DTF of 20.7±0.3 vs 23.3 ± 0.3 in the total collection) and 2016 (24.8±0.3 in ICC4363 and 27.7 ± 0.2 in the collection), and according to flowering data retrieved from field, it was one of the 29 accessions classified as early, from a total of 251 (ESM1). By contrast, when grown in SD no particular earliness for this line was observed. In any case, the restricted occurrence of these two SNP shows that they are not of broad relevance for chickpea flowering time adaptation.

Finally, the residue affected by the third SNP, Trp90, is highly conserved not only among FT but also in TFL1, BFT and MFT proteins. To investigate a possible effect of this substitution in flowering time, the accessions carrying either Trp or Leu at this position were grouped and their phenotypes compared in the three different environments (2015 LD and 2016 both LD and SD). No significant differences were found between the means of the two groups in any environment. Apart from its potential to change protein function directly, another possibility was that this polymorphism might also affect splicing, as mutations at exon/intron junctions can cause intron retention. This possibility was examined by comparing the *FTc* mRNA in wild line PI489777 (G allele) and in accessions ICC4958 and ICCL81001 (T allele). Figure 6.12 shows that only one band of the predicted size was obtained in all three accessions when amplifying *FTc* mRNA with a forward primer located in the first exon (FTc-F7b) and a combination of different reverse primer in the second, third and fourth exons. Therefore, we can conclude that the SNP is not causing the retention of the second intron or otherwise affecting splicing. There is thus no clear evidence that this particular SNP is functionally significant.

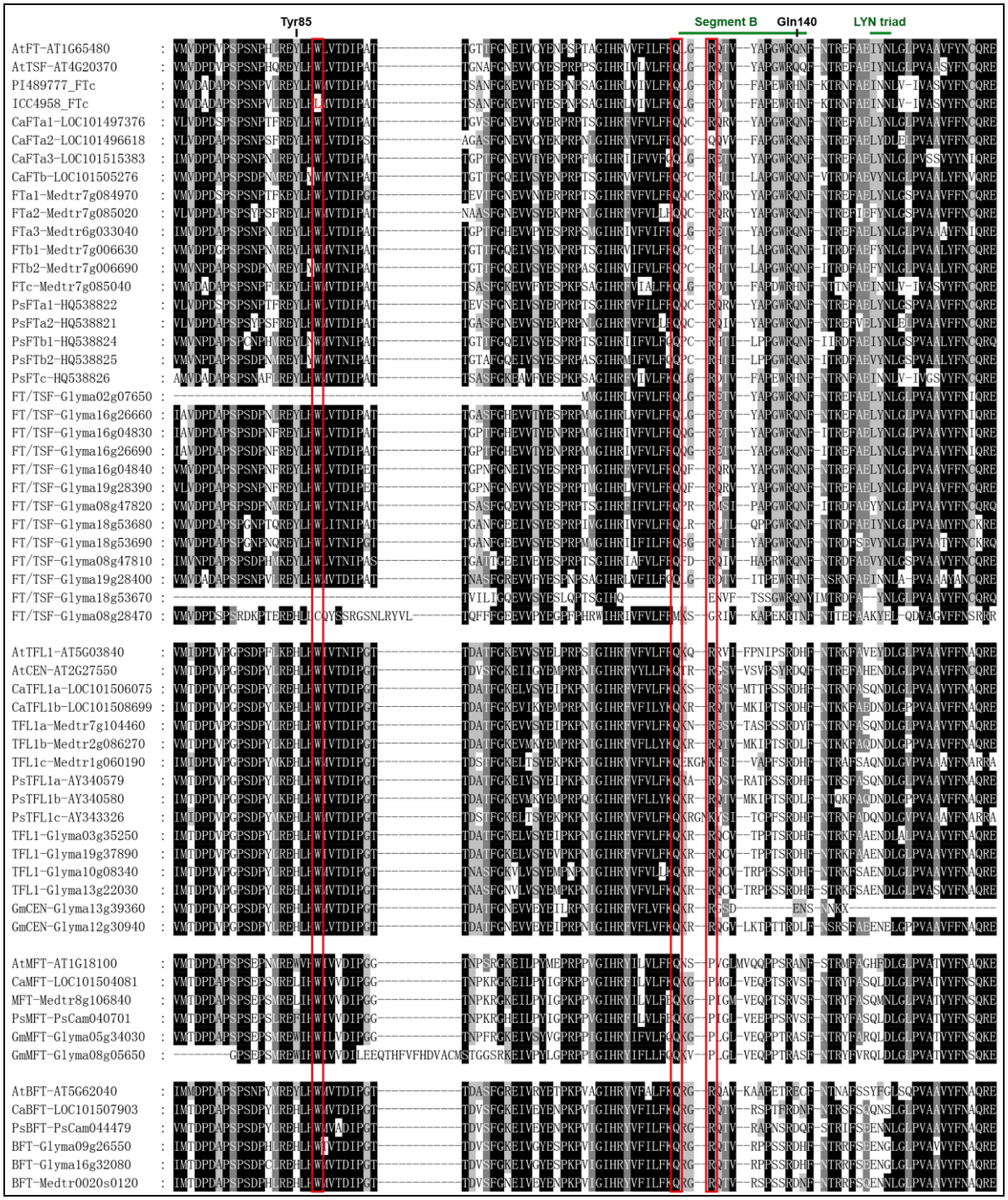


Figure 6.11 Partial alignment of PEBP proteins from Arabidopsis and various legumes showing the residues affected by SNPs identified in CaFTc. Trp88Leu, Arg130Met and Glu127Leu are indicated by red boxes. Amino acids known to be important for FT/TFL1 activity in Arabidopsis, such as Tyr85, Gln140, Segment B and the LYN triad are indicated over the alignment (Hanzawa et al. 2005; Ahn et al. 2006; Ho and Weigel 2014). (Ca = *Cicer arietinum*; Ps= *Pisum sativum*; Medtr = *Medicago truncatula*; Glyma = *Glycine max*).

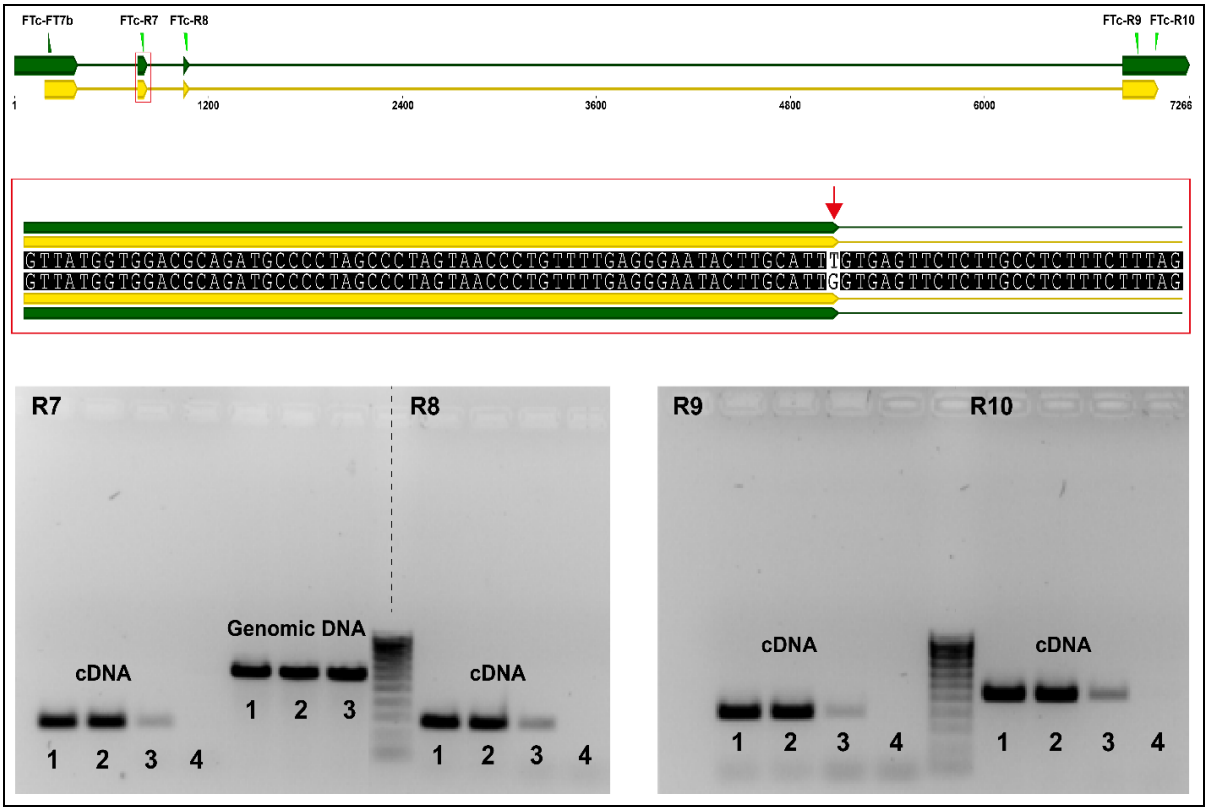


Figure 6.12 The top panel shows a diagram of the chickpea *FTc* gene with the position of some primers. mRNA indicated in green and coding sequence in yellow, with exons represented by boxes and introns by lines. The red box indicates the area expanded in the central image (second exon and beginning of third intron), with the non-synonymous SNP highlighted by a red arrow. On the bottom, PCR products obtained with the primer FTc-F7b and different reverse primers (**R7**, **R8**, **R9** and **R10**) in cDNA or genomic DNA. Lanes as follow: 1-PI489777; 2-ICC4958; 3-ICCL81001; 4-Non- template control.

6.3.5. Phylogenetic inferences

In order to examine the relatedness of *FT* cluster sequences among the different accessions, A neighbour-joining phylogenetic tree (Fig 6.14) and a Median-joining haplotype network (Fig 6.13) were constructed.

A total of 32 haplotypes were identified by PopART. Overall, the haplotype network has a star-like shape (Fig 6.13), centred around one that represent the ancestral cultivated haplotype. Except for the haplotype representing *C. reticulatum*, which is more divergent, most of the other haplotypes differ by only a few mutations and are therefore variants of the ancestral. This pattern suggests that they originate recently and is indicative of a population expansion during the recent history of the species. The neighbor-joining tree obtained **(**Fig 6.14) show similar results; few of the accession were grouped with a threshold 50% of bootstrap support specified in the analysis. In both the neighbour-joining tree and haplotype network, the

accessions show a lack of structure, as *desi* and *kabuli* accessions show no clear genetic distinction according to the *FT* genes, and only a weak tendency to group according to the geographic origin.

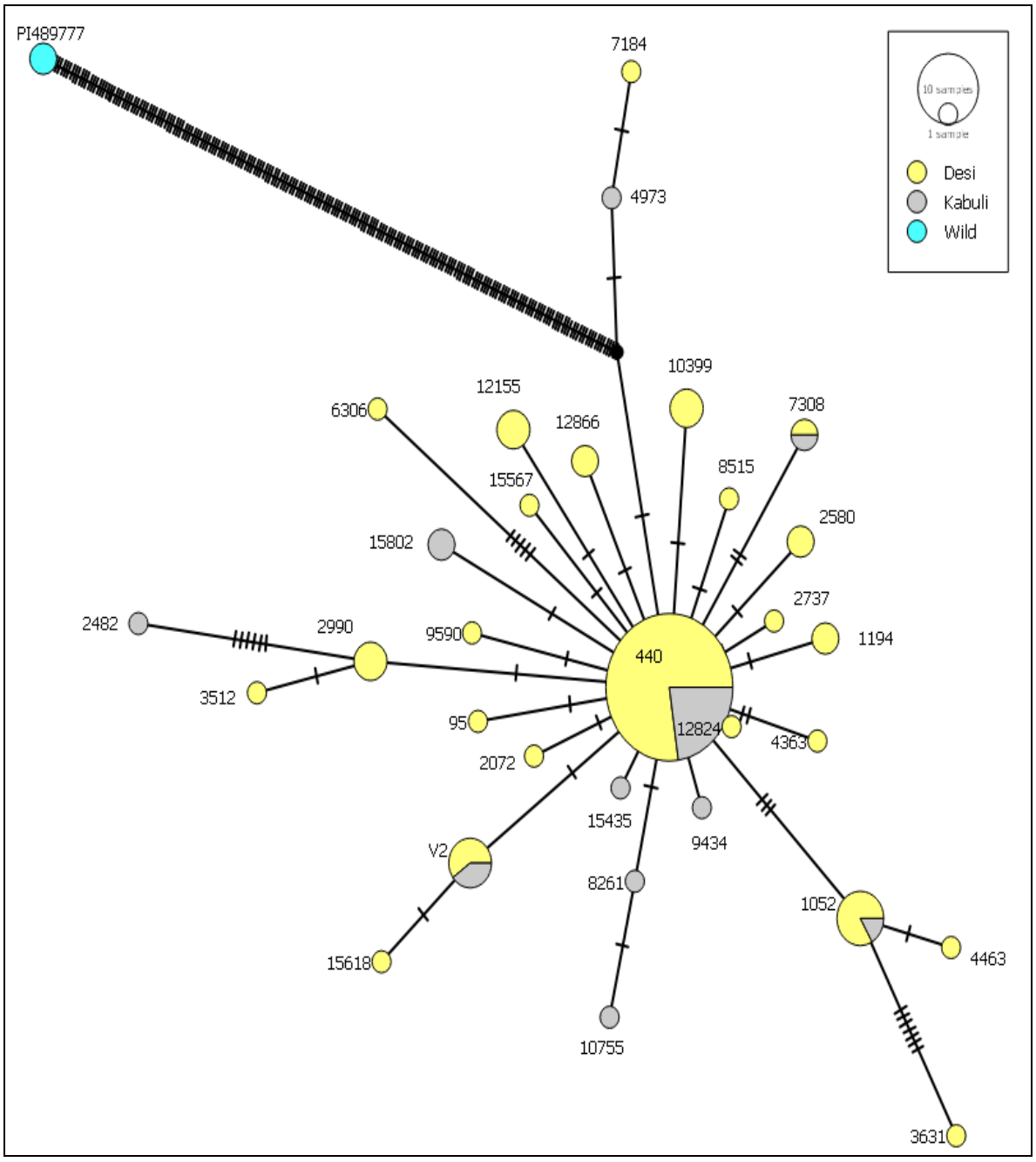


Figure 6.13 Haplotype network based on the entire *FT* cluster sequence of a chickpea collection and constructed using the median-joining algorithm. Unique haplotype groups are represented as circles connected by hash marks that indicate base pair changes between haplotypes. Numbers on haplotypes indicates a representative accession of that haplotype (excluding the "ICC-" prefix). A list of all accessions sharing each haplotype can be found in appendix 6.5.

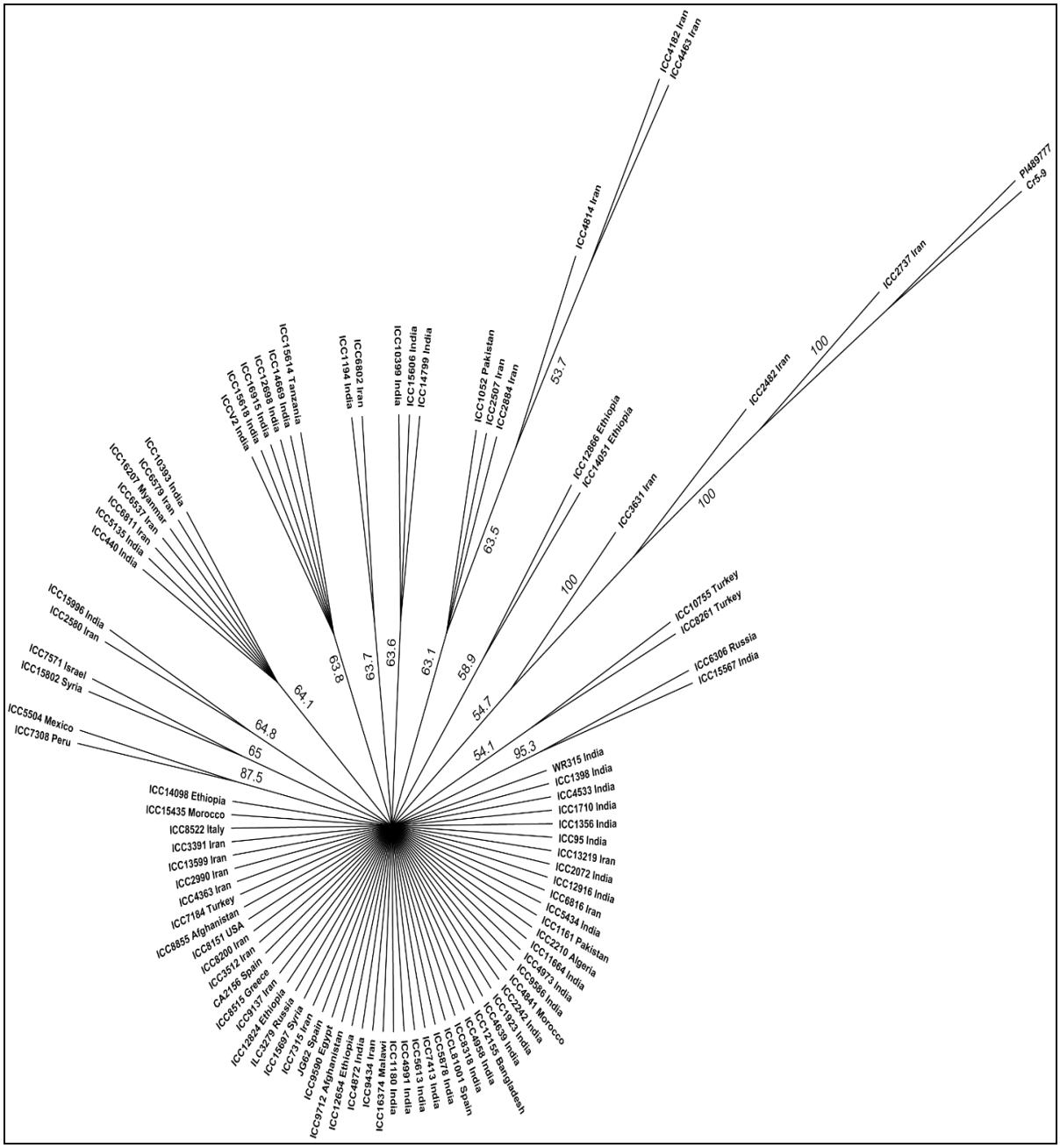


Figure 6.14 Neighbour-joining phylogenetic tree based on the alignment of 96 DNA sequences from C. reticulatum and *C. arietinum*. Sequences were aligned using ClustalW in Geneious 8 software. Numbers in branches indicates bootstrap support from 1000 replicates, with a minimum of 50% necessary to group accessions.

6.4 Discussion

Cultivated chickpea exhibits an earlier flowering phenotype compared to its wild predecessor, and this is understood to result from of selective pressure exerted since the early Bronze age when chickpea underwent a transition from winter to summer cropping (Berger et al. 2004a). In chapter 4 of this thesis, two intraspecific population were used to map the source of this earliness to a region of chromosome 3 containing a cluster of *FT* genes. The expression patterns of chickpea *FT* genes, analysed in the same chapter, is consistent with their potential role in flowering promotion, as proven for phylogenetically close legume species (Laurie et al. 2011; Hecht et al. 2011). Differences observed in their expression between the parental accessions of the crosses suggest that early flowering of the domesticated lines is associated with earlier induction and higher expression levels of these genes, suggesting the possibility of cis-regulatory changes resulting in their de-repression. Here, we characterized sequence variation across the cluster as a first step in investigating the possible causes of this elevated expression.

Interspecific polymorphisms

Not surprisingly, a large number of polymorphisms were observed between the *C. reticulatum* and *C. arietinum* parental lines in the ~55 kb region sequenced. The transcribed regions of the three genes were well conserved, with only one non-synonymous mutation found in *FTc* second exon. Although this affected a highly conserved residue, no clear evidence could be found to link this change to a difference in flowering time or FT protein function. Also, both alleles appeared to be relatively common within the subset of domesticated lines examined, suggesting that the polymorphism arose after domestication.

Most of the differences between the wild and cultivated *FT* sequences were found in the non-coding regions. The most striking sequence difference found in this study is a Ty1/copia type LTR-retrotransposon insertion in the third intron of *CaFTa*. This insertion was present in all cultivated alleles but not in the wild species, implying that the insertion may have taken place after the domestication process. However, due to the low number of wild accessions analysed here, this conclusion should be taken cautiously. Further sequencing of the cluster in a higher number of *C. reticulatum* accession is needed to confirm it. Such work is currently underway; one of the objectives of the Consultative Group for International Agricultural Research

(CGIAR) Consortium Research Program 3.5 on Grain Legumes is to assess the molecular diversity among C. reticulatum accessions held in their repositories (http://chickpealab.ucdavis.edu/index.php/goals/cgiar-consortium-research-program-35/).

The presence of transposable elements (TEs) can have a profound impact in the surrounding genes, a well-documented phenomenon that has been the subject of several recent reviews (Cui and Cao 2014; Cui and Cao 2015; Deng et al. 2016; Jern and Coffin 2008). In plants, LTR-transposons as that reported in the present study are an important source of gene evolution and, therefore, of plant variation and adaptation (Galindo-González et al. 2017); they can influence nearby genes by a wide range of diverse mechanisms, including altering splicing and polyadenylation patterns, affecting chromatin structure or by functioning as enhancers or promoters of gene transcription. The insertion of TEs in non-translated regions can also lead to aberrant splicing (Hirsch and Springer 2017), but no evidence was seen for different variants of *FTa2* mRNA between the wild and domesticated chickpea accessions.

There are several examples of how the insertion of a retrotransposon in the upstream region of a flowering-related gene can impair gene expression and affect flowering time (Bratzel and Turck 2015); In rice, a retrotransposon insertion in the upstream region of *Hd1* (a rice *CO* homolog) causes a weaker expression pattern and a decreased photoperiod sensitivity of this variant allele comparted to wild type (Yano et al. 2000; Hori et al. 2016). In soybean, a *Ty1/copia*-like retrotransposon inserted in the first intron of the *E9* locus (a soybean *FT* ortholog) attenuated its expression although no disruption of the mRNA was found (Lee and An 2015). The opposite effect, where retrotransposon insertions cause upregulation of the gene where is inserted, is also documented (Galindo-González et al. 2017). In a classic example, the increased apical dominance of maize compared to its wild ancestor teosinte is due to the enhanced expression of the gene *TEOSINTE BRANCHED 1* (*TB1*), caused by retrotransposon insertion in the 5' regulatory region (Studer et al. 2011). Similar reports can be found in the case of flowering genes; in wheat, a retrotransposon insertion in the promoter region of *VRN3* (*FT* homolog) is associated with elevated expression of this gene and thus early flowering (Yan et al. 2006). In Medicago, retrotransposon insertions in either the last intron or in the 3' region of *FTa1* resulted in gain-of-function alleles conferring elevated *FTa1* expression and early flowering (Jaudal et al. 2013). These examples provide a potential explanation for the constitutively high *FTa2* expression in all cultivated lines reported in chapter 4. Since *FTa2* is the gene were the insertion is localized, this gene would be the main affected, but changes in the chromatin and the recruitment of transcription factors needed to

start transcription could be extended to surrounding genes and affect *FTa1* in this case, due to its genomic proximity (less than 4 kbp).

Besides the transposon insertion, a high number of polymorphisms were found in the non-coding regions in the comparison of the *FT* cluster from PI489777 and ICC4958 accessions. Variation was particularly rich in the *FTa1*upstream region, where hundreds of SNPs and indels were detected between both alleles. In view of the evidence pointing to*FTa1* as the key functional gene regulating flowering time in this cluster in other legumes (Laurie et al. 2011; Putterill et al. 2013; Hecht et al. 2011) and in chickpea in particular (Chapter 4), polymorphisms in the *FTa1* promoter and introns might have particular relevance for regulation of its function. In Arabidopsis, the *FT* gene is tightly regulated by multiple transcription factors that targets specific sequences in regions around the gene and modulate its expression. Various proximal and distal regions in its promoter are crucial for proper regulation, such as two CO-responsive elements (CORE1 and 2), a distant CCAAT box (5.3 Kb upstream the ATG codon, recognized by Nuclear factor Y and involved in the stabilization of CO that is binded to CORE1/2) or a Dof-binding site for *CYCLING DOF FACTOR*, among others (Tiwari et al. 2010; Adrian et al. 2010; Cao et al. 2014; Andrés and Coupland 2012; Song et al. 2015). The distance between distal and proximal elements is determinant for its photoperiodic response; differences in the length of this region have no effect in LD. By contrast, deletions in this area result in elevated *FT* expression and early flowering under SD, while insertions have the opposite effect (Liu et al. 2014). In the same way, both the promoter and first intron contains a CArG box that is targeted by the flowering inhibitor *FLC* to repress *FT* (Helliwell et al. 2006). In legumes, variation in non-coding regions has also been linked to differences in *FT* expression. Polymorphisms in the UTR and intronic regions are associated with higher expression of an *FT5a* allele that causes early flowering in soybean (Takeshima et al. 2016). In narrow-leafed lupin, a deletion in the promoter region of an *FTc* homolog is associated with massively upregulated expression, and the loss of vernalization requirement for early (Nelson et al. 2017). Similar effect of cis-variation has also been reported in other flowering-related genes. For example, natural variation in the non-coding sequences of Arabidopsis *FLC* define five major haplotypes that display differential epigenetic silencing and expression levels of *FLC*, which in turn determine the degree of vernalisation responsiveness in accessions from different latitudes (Coustham et al. 2012; Li et al. 2014). A similar scenario has been described in *Brassica oleracea*; allelic variation at the *BoFLC.C2* locus (an *FLC* orthologue) determines the

effectiveness of the epigenetic silencing that vernalization exerts on this gene. In functional alleles, *BoFLC.C2* stays repressed after vernalization, whereas in others the expression of this gene gradually recovers pre-vernalized values after transfer to warm, causing a late phenotype. Although the exact cause is still unknown, polymorphisms in the non-coding intronic regions, rather than those found in coding sequence, are more correlated with the differential vernalization response (Irwin et al. 2016).

Summarizing, a high number of polymorphisms between the wild and cultivated alleles of the *FT* cluster were observed. Which ones among them could be involved in the observed upregulation of these genes is still to be determined. Further research will be needed, analysing potential regulatory motifs present in the wild accessions and how the polymorphism presented in this study modify them in the cultivated germplasm.

Intraspecific polymorphisms

As expected in a crop species, the number of polymorphic sites was lower at the *C. arietinum* intraspecific level. Yet, almost 100 polymorphisms were detected among the 94 accessions analysed.

Similar to results obtained in the interspecific comparison, the coding sequences were highly conserved. Besides the substitution already discussed in the previous section, two more amino acid changes were found in the FTc protein. Only one of them, the change Arg130Met found in accession ICC4363, showed potential to confer early flowering in mutagenesis studies in Arabidopsis (Ho and Weigel 2014). Interestingly this accession was one of the earliest lines in both the field and phytotron under LD. However, its earliness might be due to another factor, so in order to confirm its influence on flowering, further screening of this mutation in the large chickpea collections available at research centres is needed to see its spread through the available germplasm and test whether they present or not association with any flowering time alteration.

The non-coding region accumulated most of the polymorphism. The most significant changes, because of their size and potential association with flowering time, were two big deletions (~30 kb each) and a ~750 bp insertion in the *FTa1-FTa2* intergenic region. The deletions (*type 1* and *type 2*) could be a source of variation for flowering time, as we found that chickpea lines with *type 1* were always later to flower, whereas those with *type 2* are later in LD but seems to be earlier in SD. Due to their big size, both deletions affect the complete

genomic sequence of *FTa2* and also the *FTa1-FTa2* intergenic region. Each of these elements have the potential to cause alterations in flowering time, although *FTa2* could be considered a less likely candidate for a number of reasons; first, the loss of this gene did not obviously affect any aspect of plant development or architecture, suggesting that perhaps *FTa2* is not functional or might be highly redundant with its paralog *FTa1*. The high homology between *FTa1* and *FTa2* proteins and their tandem arrangement is shared in all galegoid legumes so far examined (Laurie et al. 2011; Hecht et al. 2011; Rajandran 2016) indicating that these genes have arisen through duplication in a common ancestor of this group. Second, Arabidopsis complementation studies realized in Medicago and pea suggest that this gene is the least effective for promotion of flowering among all *FT* examined (Laurie et al. 2011; Hecht et al. 2011).

By contrast, the region between *FTa1* and *FTa2* seems to be very important for the regulation of *FTa1* expression; in Medicago, retrotransposon insertions in the orthologous region results in elevated *FTa1* expression and early flowering as reported by Jaudal et al. (2013). A ~10 kb deletion in the same region has been associated with a major flowering QTL in lentil (Rajandran 2016). Finally, further polymorphism observed in this study also support the possible relevance of the *FTa1-a2* intergenic region as a 750bp insertion was significantly associated with early flowering when plants were grown under SD.

Which part of the intergenic region is relevant in flowering regulation is unknown, but the intergenic region in chickpea is quite small, comprising only 3.6 Kbp. Most of it is occupied by a non-coding RNA (ncRNA) spanning 2.3 kbp, according to NCBI annotations (GeneID 105851807). Interestingly, both *type 1* and *type 2* eliminate the last portion of this ncRNA. The relevance of this ncRNA is not clear, but the role of such RNAs in gene regulation have gained notoriety in recent years, uncovering an enormous potential at the transcriptional, translational, and mRNA-stability level (Kaikkonen et al. 2011; Inouye 1988; Clancy 2008; Lau and Lai 2005). They can also act as a decoy for microRNAs (miRNAs), blocking their interaction with the authentic targets (Wu et al. 2013a). Interestingly, in the model species *Brachypodium distachyon* there is evidence of miRNAs participating in the regulation of *FT* genes (Wu et al. 2013b). More research is needed to define the extent and splicing pattern of this ncRNA, to determine which parts of the region are involved in the modulation of *FTa1* expression and to investigate their mechanism(s) of action.

A final question that arise from the results is that regarding the quite unique features observed in the case of chickpea *FTa2*. Despite an apparent lack of functionality, as discussed above, the features observed in this gene are at least worth a mention. In first place, its structure of seven exons and six introns is quite different to that of canonical *FT* genes, which consist of four exons and three introns. The transcription of the three additional exons in the 5' UTR of this gene, annotated in the NCBI entry (GeneID 101496618), were confirmed here. Such structure has not been reported for any *FT* gene from legume species. Also, in the present chapter, we detected proliferation of distinct splice variants for *FTa2*. More than ten isoforms were identified from PCR using a single primer pair, and the use of different combinations could show an even higher number. Interestingly, half of these transcripts lacked the first coding exon, and would likely encode non-functional proteins; one fact that could possibly explain the apparent lack of function of this gene. There was no evidence for any alterations in the abundance of these alternative transcripts between wild and domesticated lines, suggesting that the RT1 insertion did not affect *FTa2* splicing.

The question of the possible functional significance of these transcript variants remains open. Alternative splicing (AS) is an important gene regulatory mechanism that, among other functions, allows proteome diversification and can influence gene expression (Stamm et al. 2005), and evidence suggests that around 20% of plant genes show AS (Wang and Brendel 2006). Compared with the extensive knowledge of *FT* control at the transcriptional level, our understanding of *FT* regulation at the posttranscriptional level is limited, although several cases of alternative splicing of *FT* genes have been documented; two *FT* genes have been found in the tree London plane (*Platanus acerifolia*), that yielded more than 34 different transcript depending on the tissues and developmental states and suggesting multiple regulatory roles for the different isoforms (Zhang et al. 2011). In maize, there are 15 *FT-like* genes and two of them (*ZCN18* and *ZCN26*) show mRNA splicing (Danilevskaya et al. 2008). Sunflower is another species in which *FT* alternative splicing have been also described, although this case looks more like an splicing defect resulting in intron retention rather than a regulated alternative splicing (Blackman et al. 2010). The grass species *Brachypodium distachyon* possess 2 *FT* genes, *FT1* and *FT2*, with redundant roles in induction of flowering. Two isoforms of *FT2* are formed by alternative splicing; *FT2α* (the fully functional isoform) and *FT2*β, which lacks a short section of the N-terminal domain and thus cannot form a functional florigen complex with FD and 14-3-3 proteins. FT1, FT2α and FT2β proteins can form heterodimers, but any complex containing the defective FT2β fails to promote flowering.

The *FT2*β/FT2α heterodimer ratio progressively decreases during development, resulting in a gradual increase in the florigen activity level (Qin et al. 2017). Finally, one of the *FT* orthologues in *Chrysanthemum morifolium* is spliced in five different ways depending on developmental stage, with each variant showing different flowering-promotion capacity. One of them is lacking the exon 2 (analogue to those missing the first exon in this study) and was unable to promote flowering (Mao et al. 2016).

Future research analysing the expression of the different chickpea *FTa2* isoforms in different developmental stages, tissues and environments will be necessary to fully understand their regulation and biological meaning. Due to the unknown effect of the retrotransposon insertion in the second intron in cultivated chickpea, *C. reticulatum* would be a preferable system in which to address these questions.

Genetic diversity among C. arietinum accessions

The nucleotide diversity found among *C. arietinum* accession analysed was very low, with values similar to those previously reported (Gujaria et al. 2011; Rajesh and Muehlbauer 2008; Jain et al. 2013). The low genetic diversity of chickpea is well documented, and is a consequence of the history of chickpea, which was domesticated around 10000 years ago during the Neolithic era. Since then, it suffered a series of four sequential bottlenecks during its evolution as a crop that highly constrained its genetic diversity, in contrast with other self-pollinating crops domesticated at the same time (Abbo et al. 2003a). The star-like shape obtained in the haplotype analysis is a reflection of this history and denotes a young species in expansion period. We could not detect any difference between *desi* and *kabuli* accessions based on their *FT* sequences, and the grouping of the accession correlated more with the geographical distribution, consistent with results obtained in previous reports (Penmetsa et al. 2016). The general lack of structure obtained is probably a consequence of frequent movement of alleles due to intense commercial activity for many centuries, the dispersion of the crop linked to the human colonization of the different continents and the use of few lines to develop elite varieties in breeding programs performed in the last century.

Concluding remarks.

A great number of polymorphisms were found at both the inter- and intra-specific level by sequencing the *FTa1*, *FTa2* and *FTc* genes in a collection of 96 *desi*, *kabuli* and wild chickpea accessions. As expected, these polymorphisms were more abundant in the non-

coding region. There are many examples showing that polymorphism in non-coding areas can create functional diversity, so it is possible that one or more of the identified polymorphisms between wild and cultivated chickpea could impair the function of a sequence needed for transcriptional repression of *FTa1* and possibly other regions of the cluster, resulting in a gain-of-function mutation conferring elevated *FTa1* expression and early flowering, as hypothesised by Weller and Ortega (2015). Which of these polymorphisms may be responsible for such a change in expression needs to be determined by future research, but the identification of variant sites is the necessary first step in this direction. The two major indels affecting the *FTa1*/*a2* intergenic region and the atypical *FTa2* promoter seem most promising based on recent reports from Medicago about the functional relevance of this intergenic region (Jaudal et al. 2013) and results from our research group showing association between a large *FTa1/2* intergenic deletion, elevated *FTa1* expression and early flowering in lentil (Rajandran 2016).

The results obtained here evaluating the effect of the polymorphism, however, are not conclusive, due to the small number of accessions analysed. Future research is needed to investigate the extent of the detected polymorphism in the cultivated germplasm. Once an adequate number of lines with each of the polymorphism is detected, their association with flowering time variation should be tested in a wider range of environments. Another limitation of the present study, already stated in the discussion, is that it is based the use of only two wild accessions, and it therefore remains unclear whether any polymorphisms identified here are indeed representative, so any conclusion on their potential functional significance or role in domestication will need further comparison with sequences from a larger number of diverse of *C. reticulatum* accessions, and a clearer understanding of the likely wild progenitor.

Chapter 7. Regulation of vernalization in legumes

7.1 Introduction

Vernalization is defined as the process by which prolonged exposure to cold temperatures promotes flowering (Amasino 2004). Similar to other species in the temperate galegoid clade of legumes, wild species within the genus *Cicer* possess a strong vernalization requirement, including *C. reticulatum* from which the domesticated form *C. arietinum* derives (Berger et al. 2004a; Abbo et al. 2002; Berger et al. 2005; Sharma and Upadhyaya 2015). However unlike its wild predecessor, domesticated chickpea has been widely considered to be vernalization insensitive, and it has been suggested that vernalization response was lost during the domestication process, along with other winter-adaptive traits such as low temperature tolerance during the vegetative and reproductive phases (Abbo et al. 2003a; Berger et al. 2012; Summerfield et al. 1989). The majority of wild *Cicer* species can be found in geographic areas that expose them to vernalising temperatures on an annual basis (Berger et al. 2003), whereas most domesticated chickpea is nowadays grown either as a spring crop or as a winter crop in subtropical areas and thus does not experience this low temperature exposure. With this in mind, it seems logical that relaxation of the strong vernalization requirement of *C. reticulatum* was a necessary step in the evolution of chickpea as this requirement if not met would significantly delay flowering, extend crop duration, and impose a severe yield penalty (Abbo et al. 2009; Abbo et al. 2003b).

More recent findings suggest that this assumption could be only partially true, as a subgroup within the cultivated chickpea pool was found to be vernalization-sensitive (Sharma and Upadhyaya 2015; Pinhasi van-Oss et al. 2016). However, even in these cases, the response is apparently weaker in *C. arietinum*, as the advance in flowering caused by vernalization was smaller compared to wild *Cicer* species. Although the majority of chickpea breeding programs are focussed towards the creation of short-cycle varieties with early phenology, variability in vernalization and photoperiod responsiveness could potentially be valuable in the development of cultivars for different agroclimatic regions (Pinhasi van-Oss et al. 2016), and further research is needed to investigate the diversity and geographical distribution of vernalization responsiveness within chickpea germplasm.

Further research is also needed into the molecular and physiological basis for the vernalization response; in legumes in general, and in chickpea in particular. The genetics

underlying vernalization has been intensively studied in *Arabidopsis thaliana* (Kim et al. 2009) where the vernalization response is mainly attributed to the floral repressor *FLOWERING LOCUS C* (*FLC*). This gene is the target of a complex regulatory network that ensures a drastic reduction of *FLC* expression after a period of cold ambient temperatures (Alexandre and Hennig 2008). Since the repressive effect of *FLC* on flowering is maintained by preventing expression of *FT* expression, the final outcome of vernalization in Arabidopsis is the upregulation of *FT*, with the consequent activation of the floral meristem identity genes that leads to flowering.

Since the *FLC* gene seems to be absent in many vernalization-responsive species outside the *Brassicaceae* (Alexandre and Hennig 2008), it follows that a response to vernalization has probably arisen independently on multiple occasions and the molecular sensing and signalling mechanisms are likely to be different in different plant clades. For example, in temperate cereals, a model involving 3 central genes (*VERNALIZATION1* to *3*; *VRN1*, *VRN2* and *VRN3*) has been proposed. These genes are homologs of Arabidopsis *AP1*, *CO* and *FT*, respectively (von Zitzewitz et al. 2005; Yan et al. 2004; Yan et al. 2003; Yan et al. 2006). *VRN1* controls the floral promotion by vernalization in these species; its expression is induced during a cold treatment and it is maintained at a high level after transfer to warm conditions. *VRN1* promotes flowering by decreasing the expression of flowering repressors (such as *VRN2*) while at the same time upregulating the expression of the floral promoter *VRN3 (*Oliver et al. 2013*)*. In sugar beet (*Beta vulgaris*), three genes involved in the vernalization response have also been described; *BOLTING TIME CONTROL1* (*BvBTC1*), which is a *PSEUDO RESPONSE REGULATOR* similar to Arabidopsis *PRR3* and *PRR7* genes, and two *FT* family genes *BvFT1* and *BvFT2*. Once again, the nature of their interaction is similar; *BvFT2* acts as a flowering promotor whereas *BvFT1*, despite belonging to the *FT* family, plays the repressive role whose expression decreases with cold exposure. In this case *BvBTC1* is the target of vernalization that acts upstream of both other genes, upregulating *BvFT2* and downregulating *BvFT1* (Pin et al. 2012; Pin et al. 2010).

These examples suggest that de-repression of *FT* expression following cold may be a widely conserved feature of the vernalization pathway in most plants. In legumes, three subclades of *FT* genes have been described, namely *FTa*, *FTb* and *FTc*. Variation in the number of genes belonging to each category has been found in different species, but in all they seem to be involved in the control of flowering (Hecht et al. 2011; Kong et al. 2010; Laurie et al. 2011). Despite the fact that molecular pathways regulating vernalization are just starting to be

investigated in legumes, research so far seems to point to *FT* genes being crucial. For example, in narrow-leafed lupin (*Lupinus angustifolius*), the *Ku* locus conferring dominant vernalization-insensitivity has been associated with *LanFTc1*, an *FT* belonging to the *FTc* subclade (Nelson et al. 2017). In *Medicago truncatula*, a species taxonomically much closer to chickpea, another *FT* ortholog (*MtFTa1*) was found to be the main target of vernalization (Laurie et al. 2011). These results suggest that *FTa* and *FTc* genes in particular may have an important role in the vernalization response of other legumes, and their characterization might give valuable information that would help in better understanding the basis for diversity in this important trait in a range of temperate legume crops.

Chapter 6 presented results of an analysis of sequence variation within the *FTa1-a2-c* cluster in a collection of diverse chickpea lines. Among cultivated accessions, two indel events were found that deserve special attention; the first was a large deletion (~30 kb) spanning the entire *FTa2* gene and parts of the adjacent *FTa1-a2* and *FTa2-c* intergenic regions, and the second was a ~750 bp insertion in the *FTa2* promoter. Polymorphism in the corresponding *FTa1-a2* region in lentil (*Lens culinaris*) has also been detected by our research group: a ~10kb deletion has been characterized in the lentil accession ILL2601 and associated with the early flowering phenotype of this line (Rajandran 2016).

This chapter will address three questions. First, the possible existence of a vernalization response in cultivated chickpea will be evaluated by analysing the phenotypic (flowering) response of several wild and cultivated accessions to a combination of vernalization and photoperiod regimes. Second, the potential role of chickpea *FT* genes as targets of the vernalization pathway will be examined, using a selection of accessions representative of the allelic variation found in the *FTa1-a2-c* cluster in both species. Finally, the possibility that vernalization regulation of *FT* genes may be conserved in other temperate crop legumes will be examined by performing a similar analysis in lentil (*Lens culinaris*), where a major QTL for flowering time is also located over the *FTa1-a2-c* cluster (Rajandran 2016).

7.2 Materials and methods

Plant material

Six *C. arietinum*, two *C. reticulatum* and two *L. culinaris* accessions were used in this study (Table 7.1).

Table 7.1 List of the accession used in the present study and *FTa1-FTa2* allele (for details, see Chapter 5)

Cicer arietinum	
ICCV2	32383 bp deletion
WR315	
ICC4958	750 bp insertion
ICCL81001	
ILC3279	No Indels
JG62	
Cicer reticulatum	
Cr5-9	
PI489777	
Lens culinaris	
ILL 2601	≈10kbp deletion
ILL5588	

Chickpea lines ICC4958, ICCL81001, WR315, ILC3279, PI489777 and Cr5-9 have been already described in the Materials and Methods section of Chapter 4 and no further information will be added here. JG62 is a *desi* accession catalogued as a medium duration variety, and ICCV2 is an extra-short duration *kabuli* cultivar (Kumar and Van Rheenen 2000).

Lentil accession ILL5588 is a landrace developed by the International Centre for Agricultural Research in Dry Areas (ICARDA), and is described to be mid-late flowering and photoperiod-sensitive, whereas lentil ILL2601 is a *pilosae* accession previously evaluated as very early to flower and photoperiod-insensitive (Rajandran 2016).

Growing condition and phenotyping

Seeds for vernalization (V) were scarified by removing a small piece of the seed coat with a razor blade, coated with Thiram to minimize the risk of fungal infection and left at room temperature in petri dishes, placed between two layers of paper towel damped with sterile distilled water for germination. Petri dishes were then moved to a cold room at 4^{o}C in total darkness for either 3 or 4 weeks in the case of lentil and chickpea, respectively. Seed of non-vernalized plants (Control, NV) were also scarified, Thiram-coated and germinated in petri

dishes with sterile distilled water 7 days prior to the end of vernalization treatment to ensure a homogenous developmental state of both vernalized and control plants at sowing time.

After vernalization treatment, both vernalized and unvernalized seedlings were sown as described in Chapter 2 (Plant growth conditions) and grown in controlled environment facilities at University of Tasmania in either long days (LD) or short day (SD) photoperiod. Plants under SD photoperiod received 8 hours of natural daylight and 16 hours of total darkness inside a phytotron, while those in LD received 8 hours of natural daylight extended with another 8 hours with white light (10µmolm-2s-1) from a mixture of fluorescent and incandescent sources (Hecht et al. 2007a). Temperature inside the phytotron was maintained at 16 ^{0}C.

Chickpea accessions were grown from January to September (2015 season) and December to September (2016 season). Flowering and podding time were recorded on each individual plant as the number of days from seedling emergence to the fully opening of the first flower (DTF, Days to Flower) or to first pod (DTP, Days to pod). The nodes bearing the first open flower (node of flower development, NFD) and the first pod (node of pod development, NPD) were also recorded in 2015. In 2016, the number of days to the visible appearance of an abortive flower bud (DTA, Days to Abortion) was also recorded. The two lentil lines were sown in the same facilities in June to November 2016. DTF was measured as described for chickpea but, in addition, the node of flower development (NFD) was also recorded.

RNA extraction and qPCR.

C. reticulatum accession PI489777 was grown under 3 different conditions; vernalized (V) plants were grown under both SD and LD, whereas control (NV) plants were grown only under SD. One day after vernalized seedlings were transferred to warm conditions, the whole apical bud (apical meristem plus the immature leafs surrounding it) was harvested in both SD and LD. As plants continued growing, all leaves up to the leaf 6 (LD-V plants), 12 (SD-NV plants) or 14 (SD-V) were collected separately and consecutively at the point they were considered to have fully expanded. In chickpea, the two first leaves (in developmental nodes 1 and 2) are scale leaves. These are not full foliage leaves and therefore they were not considered in the numbering of leaves used in this study. In addition, 46 days after transfer to phytotron, a subset of SD-V plants were placed again in cold (4^{o}C) for 2 weeks (revernalization). Only leaf 14 of these plants receiving revernalization was sampled, and this was done while the plant was still in cold conditions.

C. arietinum lines WR315 and ICCL81001 were grown at the same time and in the same conditions stated above for wild chickpea. Harvest points were also similar with the only differences being the total number of leaves harvested (only leaves 1 to 6 were collected in all conditions) and the absence of a second cold treatment.

In the case of lentil accessions ILL2601 and ILL5588, vernalized and control plants were grown in both SD and LD. During the 21-days of vernalization, the whole apical bud of seedling was collected at days 1, 3, 5, 7, 14 and 21 of treatment. Once transferred to warm conditions, the dissected apical buds and the uppermost fully expanded leaf tissue were harvested at days 28 and 35 in both SD and LD photoperiods (1 and 2 weeks after transference to phytotron).

In both chickpea and lentil experiments, samples were harvested at midday (12:00-13:00 h), and each sample consisted of pooled material from 2 plants, and two biological replicates were performed. RNA extraction and cDNA synthesis were performed as described in Chapter 2, section 2.3.2.

Chickpea primers used for qPCR were the same as those described in chapter 4. Only one additional primer pair was used to test the expression of the *FTa1-a2* intergenic region (*RMK*) and details are shown in Table 7.2, together with details of primers used for amplification of the lentil genes. In all cases, the expression of the genes was measured as described in Chapter 2, section 2.4.3 using *ACTIN* as housekeeping gene.

Table 7.2 Sequences of the primer used to measure lentil (Lc) and chickpea (Ca) gene expression

Gene	Primer name	Sequence	Reference/Origin
ACTIN	LcActin-Fw LcActin-Rv		(Saha and Vandemark 2013) (Saha and Vandemark 2012)
LcFTa1	LcFTa1-8F LcFTa1-7R	CCGATATTCCAGCAACTACTGA AACACGAACACGAAACGATG	this study
LcFTa2	PsFTL-B-2F PsFTL-B-4R	CGGAAATAGGAATGTTTCCAATGG AACTTGGGCTAGGTGCATCA	this study
LcFTb2	LcFTb1-3F LcFTb1-1R	GGTGAACCCTGATGCACCTA GAACGTTGTCCCAGTAGTCG	this study
LcFTc	LcFTc-F3 LcFTc-R3	GGTGCATCTGCGTCCACC ACAATTGGTTAATCGTCCAAGGG	this study
LcRMK	LcRMK-F3 LcRMK-R3	CAAATTTACCTGATACCACGGC GATTTGTCAACAGAACCGCC	this study
LcFTa3	LcFTa3-F1 LcFTa3-R1	AGTTCCAGGAATATCAGTCACC CCAAGGGTTAAGGTTGGTGG	this study
CaRMK	CaRMK-F1 CaRMK-R	ACTGTTCTGCACACAGTGGCTACC TGTGGTTTCTGATTTGGGGGAAGGG	this study

Statistical analysis

Analysis of variance (ANOVA) and t-test were performed using IBM SPSS Statistics software (v22.0). Chickpea phenotypic data from each season was analysed independently. Genotype and condition were treated as main effects and phenologycal traits as variables.

7.3 Results

7.3.1 Vernalization response in chickpea

To investigate the vernalization response of wild and domesticated chickpea, six *C. arietinum* and two *C. reticulatum* accessions were grown with or without vernalization treatment under either long day (LD) or short day (SD) conditions in the phytotron, in two separate experiments sown in January and December 2015. The use of the phytotron facility ensured close control of photoperiod between trials in separate years, but day temperature was poorly controlled, introducing a potential source of additional variability between the two trials. An additional difference was introduced by how the plants were handled. During 2015, plants were debranched as they grew. This has been common practice in flowering studies on other temperate crop legumes such as pea and lentil and attempts to minimize the variability in flowering time that may be introduced by differences in growth habit. In addition, it simplifies the recording of the node of the first flower and pod and, on the other hand, provides a clear visual comparison of the different accessions and the effects of photoperiod and vernalization on node number and plant height in mature plants (Fig 7.1, A). During 2016, however, plants were left intact, as shown in Fig 7.1 (B).

Although flowering time observations are often recorded as days to the opening of the first flower (DTF), it is important to mention that the node bearing the first open flower is not always the node at which the first floral structure was initiated. A common, well documented feature of chickpea development is the appearance of a variable number of nodes bearing aborted or undeveloped flowers, prior to the formation of a fully developed open flower (refer to appendix 7.1). Adverse environments such as cold and drought stress, or short photoperiods increase the appearance of flower abortions, especially in the case of late flowering genotypes (Fang et al. 2010; Nayyar et al. 2005; Zaiter and Barakat 1995). A similar behaviour has been described in the phylogenetically related legume species lentil (Rajandran 2016) and pea (Berry and Aitken 1979).

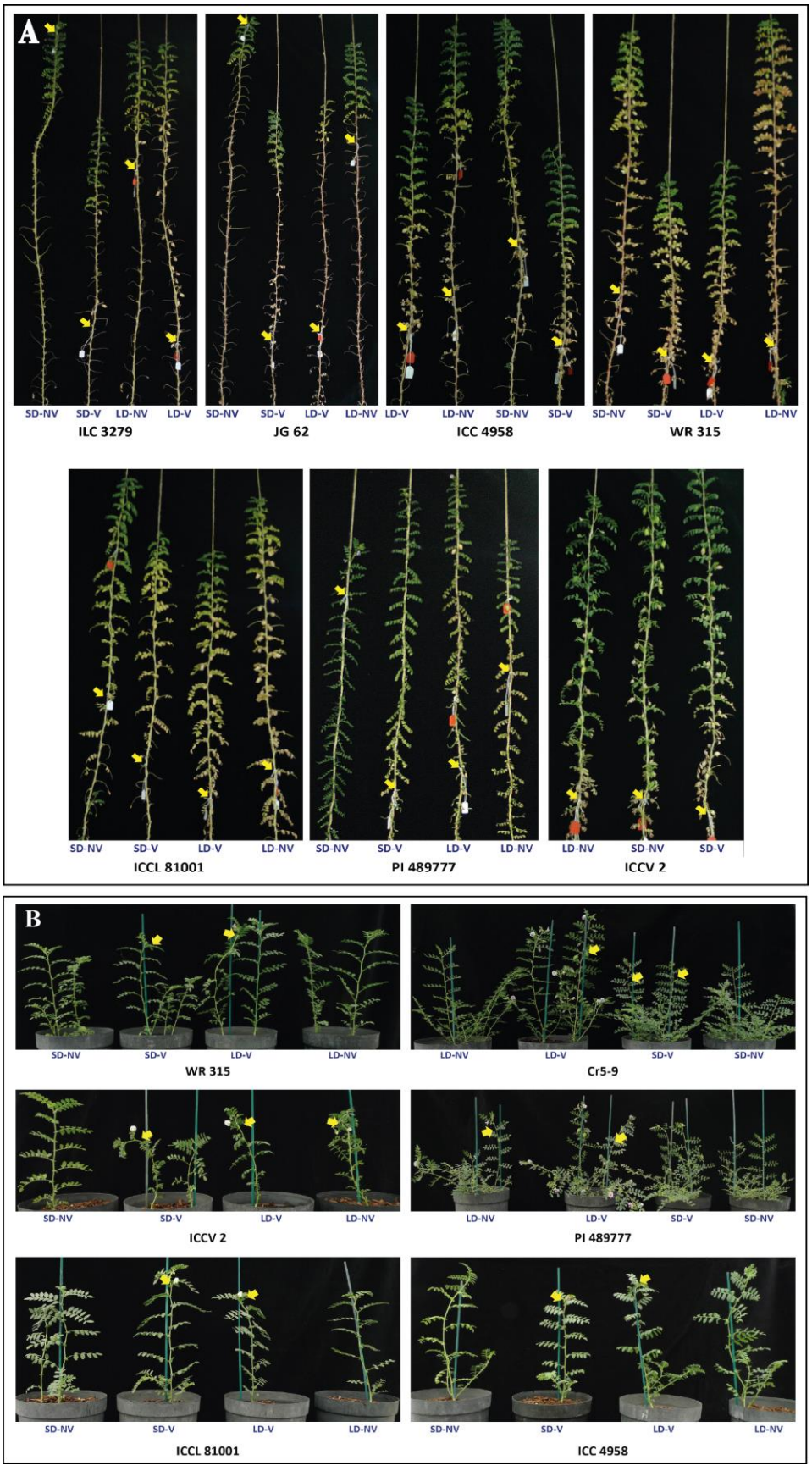


Figure 7.1 Flowering phenotype of eight chickpea accessions grown in controlled environment facilities during 2015 (**A**) and 2016 (**B**) under two different photoperiods (SD = Short day, LD = Long days) and either with (V) or without (NV) vernalization treatment. Yellow arrows point to the node of first open flower.

7.3.1.1 Effect of vernalization in chickpea flowering

A summary of the phenotypic values obtained for DTF, DTA and NFD in each genotype and season can be found in table 7.3. In general, the flowering response of the eight chickpea accessions show two clear classes (Fig 7.2); a group of four accessions (ICCV2, ICCL81001, ICC4958 and WR315) were earlier to flower in any conditions compared to the remaining four (ILC3279, JG62 and the two wild lines PI489777 and Cr5-9). From the rest of this chapter, these two groups will be referred as "early" and "late" groups.

In all eight accessions, flowering time was significantly advanced by vernalization. This result was consistent in both years (refer to appendices 7.2 and 7.3 for significance levels in 2015 and 2016 seasons, respectively). The vernalization effect was more evident in plants grown under SD (Fig 7.2), although for the early group the response was smaller due to their inherently earlier flowering phenotype. The node of first flower was also lower in all accessions in response to vernalization.

However, under LD, two contrasting vernalization responses were observed between the two groups of accessions; the early group are relatively insensitive to vernalization, whereas the late group showed a strong effect of vernalization on flowering time, similar to that seen under SD. Surprisingly, the vernalization response has a similar magnitude among cultivated and wild genotypes in this second group. Taken together, these results suggest that vernalization response exists not only in the wild *C. reticulatum* but also in all *C. arietinum* accessions. Among cultivated accessions, a subgroup behaves as vernalization-insensitive but only under flowering-inductive long photoperiods.

Another obvious effect of the vernalization treatment was a reduction in the number of aborted flowers. Long photoperiod had a similar effect (Table 7.3). In general, the strongest inductive treatments (i.e. the combination of long photoperiod and vernalization treatment) resulted in the most effective conversion of buds into open flowers best way to reduce the number of aborted buds. As expected, wild chickpea and late accessions of cultivated chickpea (JG62 and ILC3279) showed a higher prevalence of aborted flowers. By comparing DTA and DTF, it is clear that most accessions were competent to flower earlier than they actually did, but there is something preventing the development of floral structures.

Table 7.3 Sample size (N), mean (μ) and standard deviation (S) values obtained for days from emergency to first open flower (DTF, Days to flower) or to first aborted bud (DTA, Days to Abortion) in eight chickpea accession and two years, grown under a combination of long days (LD) or short days (SD) and with (V) or without (NV) vernalization treatment.

Year		2015						2016					
Trait			DTF		DTA	NFD			DTF		DTA		
Genotype		N	μ	S	N	μ	S	N	μ	S	N	μ	S
ICC4958	LDNV	10	38.4	6.4	3	17	2.3	6	31.2	3	0	-	-
	LDV	14	33.4	3.5	2	14.7	1.5	4	31.8	5.1	2	28.3	2.2
	SDNV	8	59.5	5.9	5	29.4	5.1	4	87	12.3	3	73	14.3
	SDV	11	35.2	4.3	4	16.3	2	4	56.3	20.5	3	39.3	9.2
ICCL81001	LDNV	8	34	2.4	0	15.4	0.9	15	30.1	3	0	-	-
	LDV	7	32.7	1.6	0	13.3	1.7	12	28	2.4	0	-	-
	SDNV	5	58.8	8.5	5	26.6	2.2	18	49.8	12	15	35	3.5
	SDV	2	32	0	1	15.5	0.7	14	32.2	2.8	1	32.1	2.5
ICCV2	LDNV	7	28.3	2.1	0	15.5	7.1	12	27.7	1.5	1	27.4	1.5
	LDV	14	25.9	3.1	0	13.1	1.4	14	25.9	1.9	1	25.8	1.8
	SDNV	5	39.8	5.9	3	17.8	4	10	39.8	10.7	6	35.4	6.3
	SDV	13	25.2	2.1	1	12.5	1.5	19	29.7	1.7	7	28.6	1.8
ILC3279	LDNV	5	83	15.9	1	40.2	8.2	10	60	8.4	10	38.7	5.7
	LDV	13	37.7	3.9	3	16.5	2.4	5	40.4	9.1	2	38	6.2
	SDNV	4	148	-[a]	-[a]	-[a]	-[a]	7	171	29.3	7	108.6	41.9
	SDV	8	42.3	4.2	7	18.9	2.3	5	65	16.9	5	42.6	6.1
JG62	LDNV	7	123.7	32.3	6	67.3	17.6	17	80.9	33.7	8	71.5	36.6
	LDV	11	40.6	6.4	2	18.3	3.3	15	33	3	1	32.9	2.8
	SDNV	9	148	-[a]	-[a]	-[a]	-[a]	18	163.3	20.9	18	145.4	18.5
	SDV	11	49.1	5.8	10	23.1	2.4	19	49.2	7.1	19	45.4	7.2
PI489777	LDNV	7	71.4	6.1	6	30.9	1.5	20	49.4	18.6	10	37.7	4.1
	LDV	9	38.8	4	0	15.1	1.1	19	32.5	2.1	6	31.5	2.4
	SDNV	10	131.4	10.8	7	49.5	3.5	17	168	10.6	17	123.9	18.8
	SDV	8	48.4	4.6	2	17.6	1.1	20	42.8	9.2	14	35.2	1.4
Cr5-9	LDNV	5	68.6	6.1	1	31.6	3.4	16	75.1	13.5	16	40.6	5.5
	LDV	1	39	-	0	17	-	19	31	1.8	1	30.7	1.4
	SDNV	3	131.3	12.7	2	61.7	3.1	18	171.4	11.1	18	127	17.2
	SDV	2	45	1.4	0	19	0	18	38.2	6.3	5	35.4	2
WR315	LDNV	9	36	2.2	0	17.4	1.4	17	32.2	2.2	0	-	-
	LDV	12	33.4	1.6	0	15.3	0.9	18	29.1	3.6	0	-	-
	SDNV	6	47.5	1.4	5	23.2	0.6	18	49	11.4	5	47.9	10.6
	SDV	13	35.8	1.9	0	17.1	0.9	19	36.6	3.4	0	-	-

[a.] Plants were unable to flower under this conditions during the scoring period. They were assigned a flowering time value of 148 days, coinciding with the end of the scoring period.

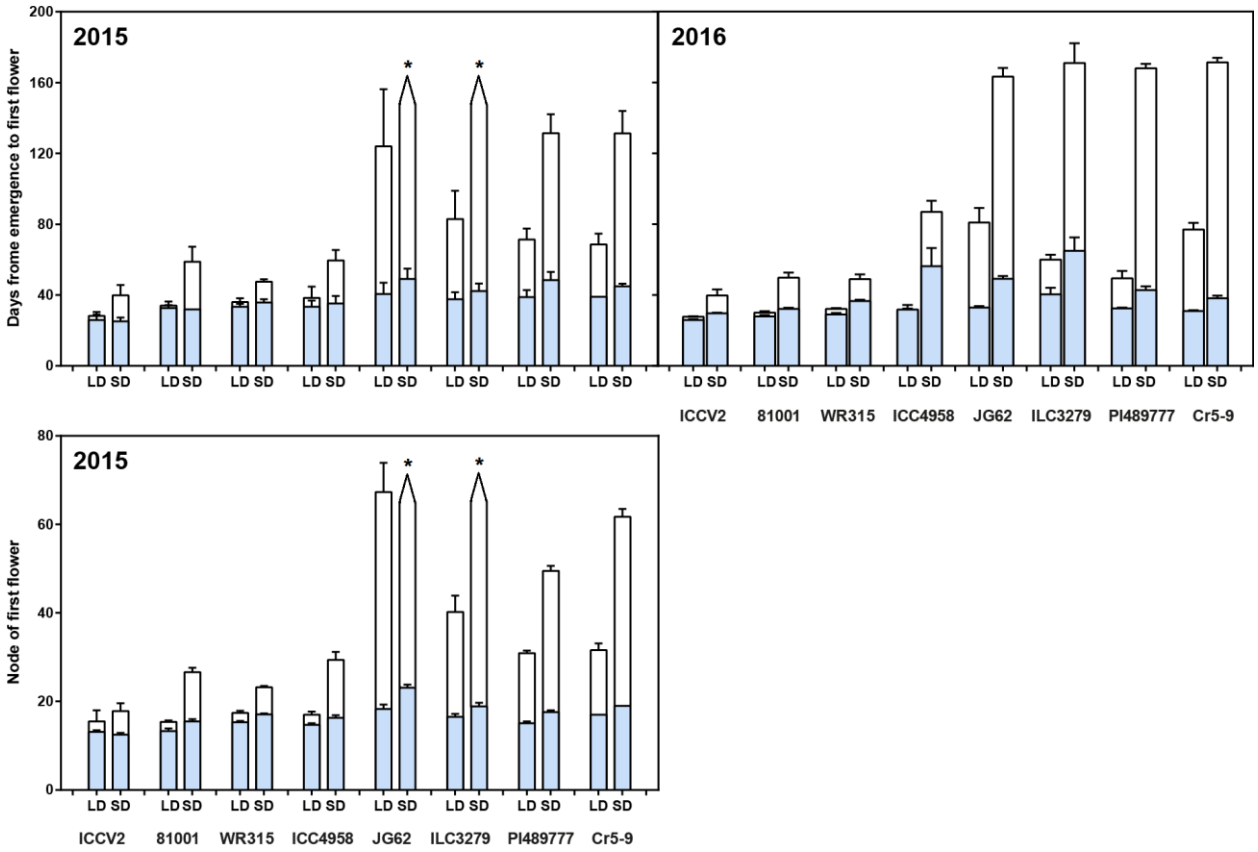


Figure 7.2 Phenotypic variation for days from emergence to first flower (2015 and 2016) and node of first flower (2015) in 8 chickpea accessions grown under long (LD) or short (SD) photoperiod and with (blue bars) or without (white bars) vernalization treatment. Arrow-shaped columns with an asterisk in the top indicate genotypes that did not flower during the scoring period.

7.3.1.2 Effect of vernalization on chickpea podding

Table 7.4 summarizes the values for days to pod (DTP) and days between the opening of the first flower to the formation of the first pod (DFTP, Days from flower to pod) obtained in each year for the eight chickpea accessions analyzed. As expected, the "early" and "late" flowering groups described in the previous section showed a similar phenology regarding podding time (Fig 7.3, A and B).

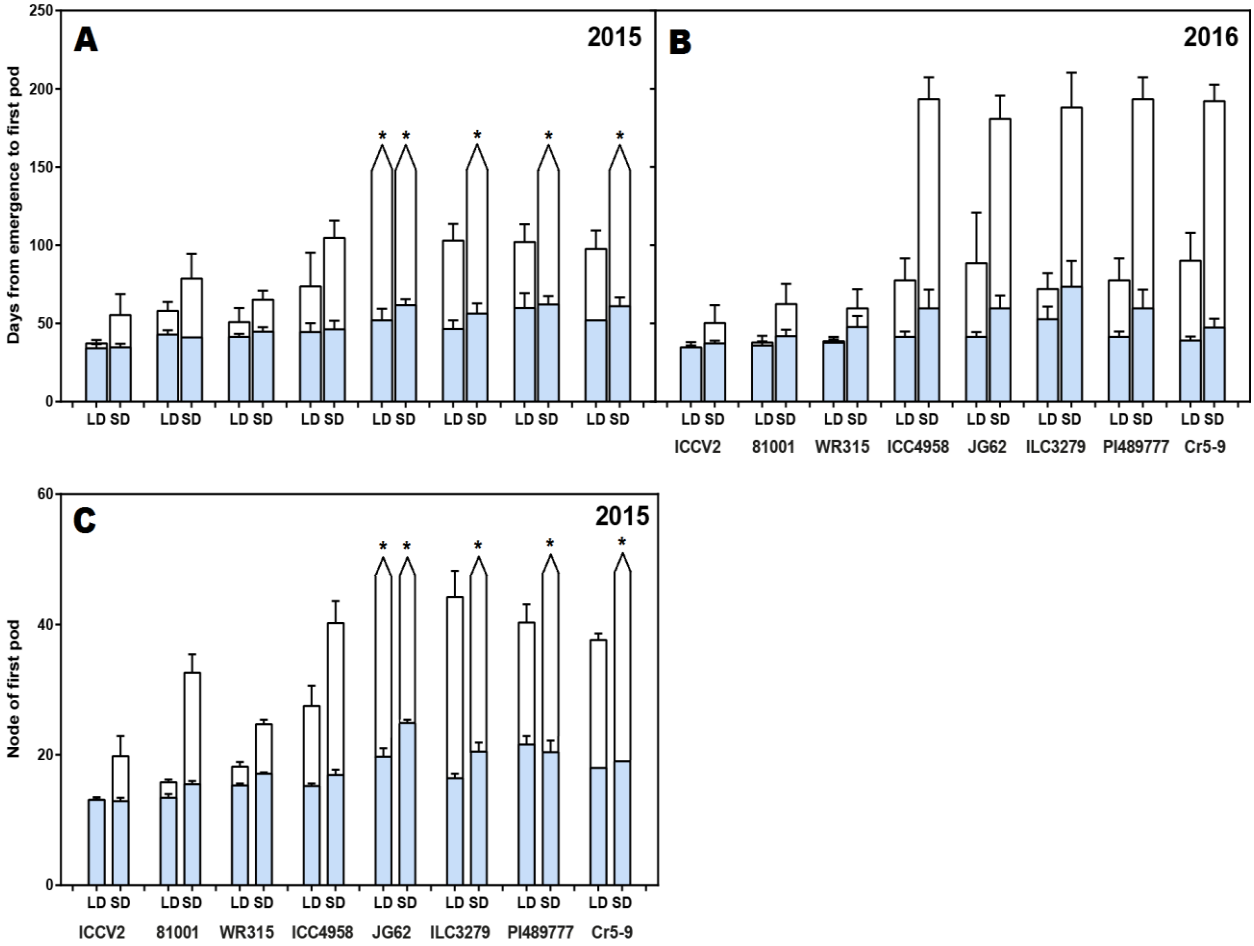


Figure 7.3 Phenotypic variation for days from emergence to first pod in 2015 (**A**) and 2016 (**B**) and node of first pod in 2015 (**C**) in 8 chickpea accessions grown under long (LD) or short (SD) photoperiod and with (blue bars) or without (white bars) vernalization treatment. Arrow-shaped columns with an asterisk above represent genotypes that did not set pods during the scoring period.

Table 7.4 Sample size (N), mean (μ) and standard deviation (S) values obtained for days from emergency to first pod (DTP, Days to Pod), node of first pod (NFP) and days from first flower to first pod (DFTP, Days from flower to pod) in eight chickpea accession and two years, grown under a combination of long days (LD) or short days (SD) and with (V) or without (NV) vernalization treatment.

Genotype	Year / Trait	2015							2016				
			DTP		NFP		DFTP			DTP		DFTP	
		N	μ	S	μ	S	μ	S	N	μ	S	μ	S
ICC4958	LDNV	10	73.7	21.5	27.5	9.9	35.3	22.4	6	42.7	6.3	11.5	7.0
	LDV	14	44.5	5.6	15.2	1.6	11.1	6.6	4	42.8	6.8	11	8.5
	SDNV	8	104.6	11.1	40.2	7.5	45.1	12.6	4	99.3	10.5	12.3	16.2
	SDV	11	46.2	5.5	16.9	2.5	11	7.0	4	73.8	18.1	17.5	27.3
ICCL81001	LDNV	8	48	5.8	15.8	1.2	14	6.3	15	37.7	4.3	7.6	5.2
	LDV	7	42.9	2.7	13.4	1.6	10.2	3.1	12	35.8	2.7	7.8	3.6
	SDNV	5	78.6	15.9	32.6	6.3	19.8	18.0	18	62.4	12.8	12.6	17.5
	SDV	2	41	0	15.5	0.7	9	0.0	14	41.8	4.2	9.6	5.0
ICCV2	LDNV	7	37.3	2.1	13.1	0.9	9	3.0	12	34.4	1.4	6.7	2.1
	LDV	14	34.1	2.7	13.1	1.6	8.2	4.1	14	34.6	3.5	8.7	4.0
	SDNV	5	55.4	13.3	19.8	6.9	15.6	14.5	10	50.3	11.4	10.5	15.6
	SDV	13	34.7	2.3	12.9	1.8	9.5	3.1	19	37.1	1.9	7.4	2.5
ILC3279	LDNV	5	103	10.7	44.2	9	20	19.2	10	72	10.1	12	13.1
	LDV	13	46.4	5.5	16.4	2.4	8.7	6.7	5	52.6	8.2	12.2	12.2
	SDNV	4	-[a]	-[a]	-[a]	-[a]	-[a]	-[a]	7	188	22.2	17	36.8
	SDV	8	56.3	6.5	20.5	4	14	7.7	5	73.4	16.6	8.4	23.7
JG62	LDNV	7	-[a]	-[a]	-[a]	-[a]	-[a]	-[a]	17	88.5	32.3	7.6	46.7
	LDV	11	52	7.4	19.7	4.2	11.4	9.8	15	41.3	3.2	8.3	4.4
	SDNV	9	-[a]	-[a]	-[a]	-[a]	-[a]	-[a]	18	180.7	15	17.4	25.7
	SDV	11	61.7	3.8	24.9	1.8	12.6	6.9	19	59.7	8.1	10.5	10.8
PI489777	LDNV	7	101.7	11.5	40.3	6.9	30.3	13.0	20	77.5	14	28.1	23.3
	LDV	9	59.9	9.5	21.6	3.9	21.1	10.3	19	41.3	3.5	8.8	4.1
	SDNV	10	-[a]	-[a]	-[a]	-[a]	-[a]	-[a]	17	193.3	14	25.3	17.6
	SDV	8	62.1	5.4	20.4	1.8	13.7	7.1	20	59.7	11.9	16.9	15.0
Cr5-9	LDNV	5	97.6	11.8	37.6	2.3	29	13.3	16	87.4	14.3	12.3	19.7
	LDV	1	52	-	18	-	13	-	19	39.1	2.5	8.1	3.1
	SDNV	3	-[a]	-[a]	-[a]	-[a]	-[a]	-[a]	18	192.1	10.5	20.7	15.3
	SDV	2	61	5.7	19	0	16	5.9	18	47.3	5.7	9.1	8.5
WR315	LDNV	9	50.8	9.1	18.2	2	14.8	9.4	17	38.6	2.7	6.4	3.5
	LDV	12	41.3	2.1	15.3	0.9	7.9	2.6	18	37.6	2.2	8.5	4.2
	SDNV	6	65.2	5.7	24.7	1.6	17.7	5.9	18	59.6	12.3	10.6	16.8
	SDV	13	44.7	3	17.1	0.9	8.9	3.6	19	47.7	7.1	11.1	7.9

[a.] Plants did not set pods during the scoring period.

Similarly to results described in the case of flowering time, vernalization produced significant advance in DTP in all accessions from both *C. arietinum* and *C. reticulatum* species in both year analysed (refer to appendices 7.4 and 7.5 for significance levels in 2015 and 2016 seasons, respectively). This effect was also reflected on the node of the first pod (NFP), scored during 2015 (Fig 7.3, C). Podding advancement was more evident in plants grown under short photoperiod, whereas under long photoperiod the bimodal behaviour described for flowering was again observed; the early group of accessions show no differences in mean DTP, while the time of first pod is reduced up to 75 % in the case of accession in the late group.

In order to test whether this advance in podding onset was a consequence of the early flowering conferred by vernalization or it was an independent effect of cold treatment, DFTP was calculated in the eight lines in both seasons (Fig 7.4). A high degree of variation was found among accessions and seasons; in general, DFTP values were lower in 2016, which could be explained by a higher mean temperature registered during this year; several sites in Tasmania (including Hobart) had their highest annual mean temperature on record, and the months between January to August (coinciding with the scoring period) were the ones that accumulated most of the differences, according to the Bureau of Meteorology of Australian Government (http://www.bom.gov.au).

Among accessions, a differential interspecific response of DFTP to vernalization was detected: the two wild lines show a clear reduction in DFTP that was consistent across conditions and years, whereas among cultivated accessions results were more variable. In 2015, vernalization treatment decreased DFTP in all *C. arietinum* lines, but in 2016 the cold treatment had (especially under LD) the opposite effect. This could reflect differences in the regulatory pathways controlling pod setting between the two species.

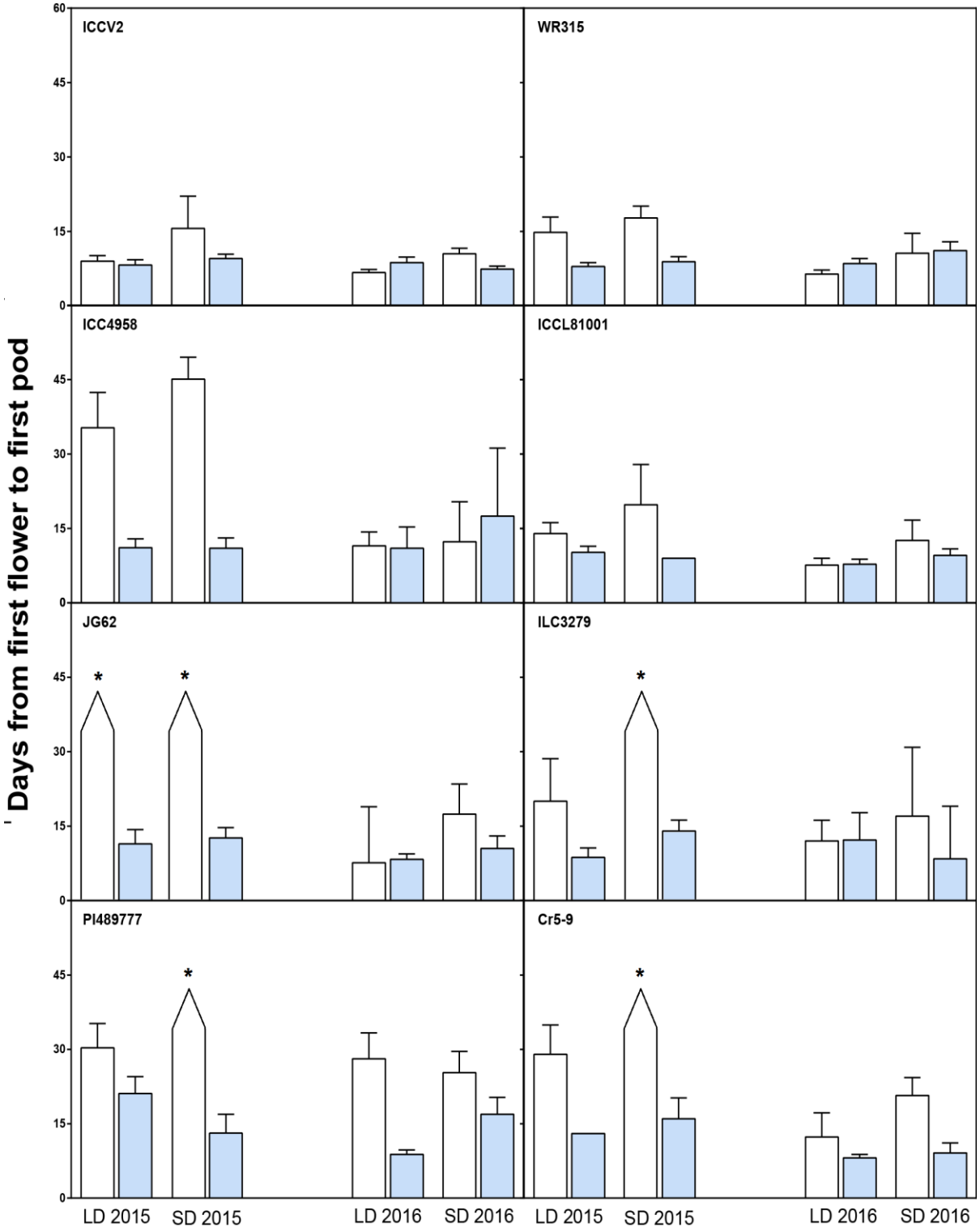


Figure 7.4 Number of days from the opening of the first flower to the first visible pod (DFTP) in eight chickpea accessions grown under long days (LD) or short days (SD) during 2015 and 2016 seasons. White columns represent DFTP values in control plants and blue columns in vernalized plants. Arrow-shaped columns with an asterisk in the top indicates plants that did not flower or set pods within the scoring period.

In summary, we found that vernalization accelerates flowering time in all chickpea accessions analysed in this study. Vernalization and photoperiod seems to share some elements in their signalling pathway. As a consequence of this overlapping, vernalization response is masked in early flowering accessions grown under inductive daylength but is evident in late flowering accessions and in all cultivated germplasm under non-inductive conditions. An advancement in the apparition of the first pod was also observed in vernalized plants, but while in *C. reticulatum* this represent another independent effect of cold treatment, in the case of cultivated lines this could be a by-product of early flowering, although further work is needed to confirm this result.

7.3.1.3 Effect of vernalization on expression of chickpea FT genes

The effect of vernalization on the expression of *FT* genes was tested in leaf tissue from three chickpea accessions, PI489777, WR315, and ICCL81001 grown under either LD or SD. The wild PI489777 accession was used as a reference, to represent a "wild-type" unmodified system where flowering pathways should be functional. The two other domesticated accessions are representative of the two major sequence variants at the *FTa1-a2* locus that were described in Chapter 6. Relative to PI489777, WR315 carries a complete deletion of *CaFTa2* gene and part of the *CaFTa1-a2* intergenic region, while ICCL81001 has a 750 bp insertion in the promoter region of the same gene.

Figure 7.5 illustrates the design for this experiment. Control (unvernalized) plants were grown only in SD, conditions in which expression of *FT* genes is minimized, as reported in Chapter 4. Vernalized plants were transferred to warm conditions under SD, to examine the effect of vernalization. Another group of vernalized plants were instead transferred to LD, in order to detect any additional effect of photoperiod. At each point in the time-course except harvest point 0, fully expanded leaves 1 to 6 (in LD grown plants) or 1 to 14 (in SD) were harvested in all treated and control plants. At harvest point 0 (one day after transfer to warm) plants had not yet formed a fully expanded leaf, so in this case, the entire apical bud was harvested.

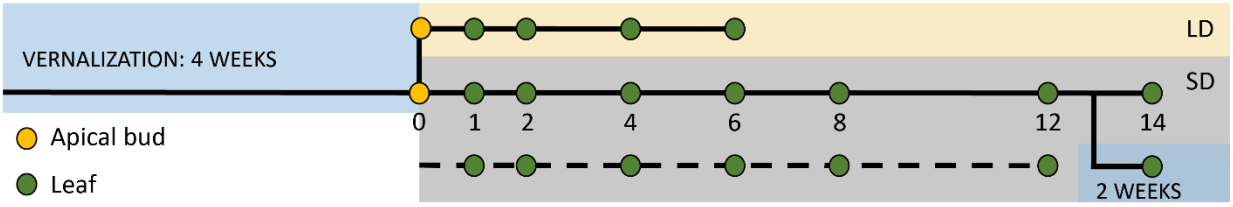


Figure 7.5 Experimental design scheme showing harvest point (filled circles) and tissues collected at each point. In the chickpea experiment, numbers indicate leaf number and in the lentil case they represent days. Continuous lines represent vernalized plants while dashed lines non-vernalized plants. SD = Short day; LD = Long day.

The expression of four of the five *FT* genes (*FTa1*, *FTa2*, *FTa3*, *FTb)* were measured Previous results in chickpea (see chapter 4) and other temperate legume species indicate that expression of *FTc* is restricted to the shoot apex, so its expression was not determined in this experiment. It has also previously been observed that regions within the *FTa1-a2* intergenic region are expressed in a cold-responsive manner in several temperate legume species (D Bond and R Macknight. pers comm.), and according to annotations from NCBI based on RNAseq data, the *FTa1-FTa2* intergenic region is also expressed in chickpea (GeneID 105851807). Therefore, primers were also designed to monitor expression of this region (it will be referred as *RMK* through this section and in subsequent discussion).

Figure 7.6 shows that in all three accessions, expression of *FTb* occurs only under long photoperiods and is not affected by vernalization. This suggests that FTb may have an exclusive role in mediating the response to photoperiod and is not directly involved in the response to vernalization. Expression levels in the two *arietinum* accessions were several-fold higher than in PI489777, but all three accessions showed a very similar pattern of FTb induction.

In contrast, clear effects of vernalization on the expression of the *FTa* genes were apparent. In the "reference" wild accession PI489777 a strong induction of *CrFTa1*, *CrFTa2* and *CrFTa3* expression was observed in vernalized plants after transfer to warm conditions, while expression remained very low in non-vernalized control plants. The induction of *FTa3* occurs earlier, starting as soon as plants are transferred to warm conditions, whereas in the case of *FTa1* and *FTa2* induction is delayed until leaf 4 (*FTa2* and *FTa1* in LD) or 8 (*FTa1* in SD). The three *FTa* gene also appeared to differ in their response to LD. Neither *CrFTa2* nor *CrFTa3* expression showed a further increase in response to long days, whereas *CrFTa1* showed a strong additional response. The maximum *FTa1* expression levels were similar

under both photoperiods, but under LD the peak of expression appeared to be shifted towards earlier leaves, peaking in leaf 8 under SD but in leaf 4 under LD. This displacement of the *CrFTa1* peak expression under LD coincides with the peak in the *CrFTb* expression. Finally, expression of the putative ncRNA *CrRMK* follows a similar pattern to *CrFTa*, as it shows qualitative induction by vernalization and is to some extent further increased in response to LD. However, unlike *CrFTa1*, which did not become significantly elevated above control (unvernalized) levels until leaf 4 (LD) or 8 (SD), the expression of *CrRMK* was high even at the earliest time point (one day after transfer from cold treatment to warm condition), suggesting that its expression might actually be induced during the cold treatment. To further investigate this possible cold induction, plants that had been maintained in SD for 46 days after vernalization, were moved back to 4^{o}C for 2 weeks for revernalization (RV) treatment. In these plants, only leaf 14 was harvested for expression analysis. Interestingly, leaves from revernalized plants show a much higher expression of *CrRMK* compared to those from equivalent plants that had not been revernalized supporting the idea that this region could be directly induced by cold exposure.

In the *C. arietinum* accessions, *FTa* expression patterns showed both similarities and differences in comparison to the wild line. In general, expression of *CaFTa3* followed very similar regulation patterns across all three accessions. However, genes in the *FTa1* region showed clear differences in the two *arietinum* lines relative to PI489777. In particular, *FTa1* itself was significantly expressed in leaves 4 and 6 even without vernalization, whereas expression is minimal in these leaves in unvernalized PI489777 plants. Also, the level of *FTa1* expression in the *arietinum* accessions did not change in response to vernalization, suggesting that the cold requirement for induction of these genes is not present in these accessions. However, both lines showed an increase in expression of *FTa1* in response to photoperiod (LD) similar to that seen in PI489777, and attained a similar expression level in leaf 4. As in PI489777, the *CaFTa1* expression pattern in LD closely followed that of *CaFTb*.

As described in Chapter 6, WR315 carries a deletion spanning the entire *FTa2* gene, and consequently no expression was detected in this accession. In ICCL81001 expression of *FTa2* was high at all time points in unvernalized plants and did not show any effect of vernalization or any additional effect of LD.

Like in the wild line, expression of *RMK* in both *arietinum* accessions showed a similar regulation pattern to *FTa1*, being responsive to both vernalization and photoperiod. Another

similarity is that the expression levels of *RMK* in vernalized plants were already high at the first harvest point but expression was not evident in control plants, suggesting that the cold induction pattern observed in *C. reticulatum* is conserved in *C. arietinum*. However, unlike the case of wild chickpea, the expression of *RMK* does not stay low in unvernalized *arietinum* plants, but it is upregulated after leaf 2 and reaches a similar expression levels of vernalized plants by leaf 4. Interestingly, this peak coincides with that of *FTa1* in control plants. Since *RMK* induction is independent of environmental cues in those conditions (SD-NV plants), it is likely an endogenous upregulation due to changes in the regulatory sequences modulating this region. It is important to remember that both *C. arietinum* accessions present alterations in these region, as described in the plant material section.

Taken together, the *RMK* cold-induced behaviour, the intrinsic upregulation of these gene in absence of any environmental input and its co-occurrence with the *FTa1* peak in control plants suggests that the *FTa1-FTa2* intergenic region could also be involved in the above-mentioned *FTa1* upregulation and loss of cold requirement.

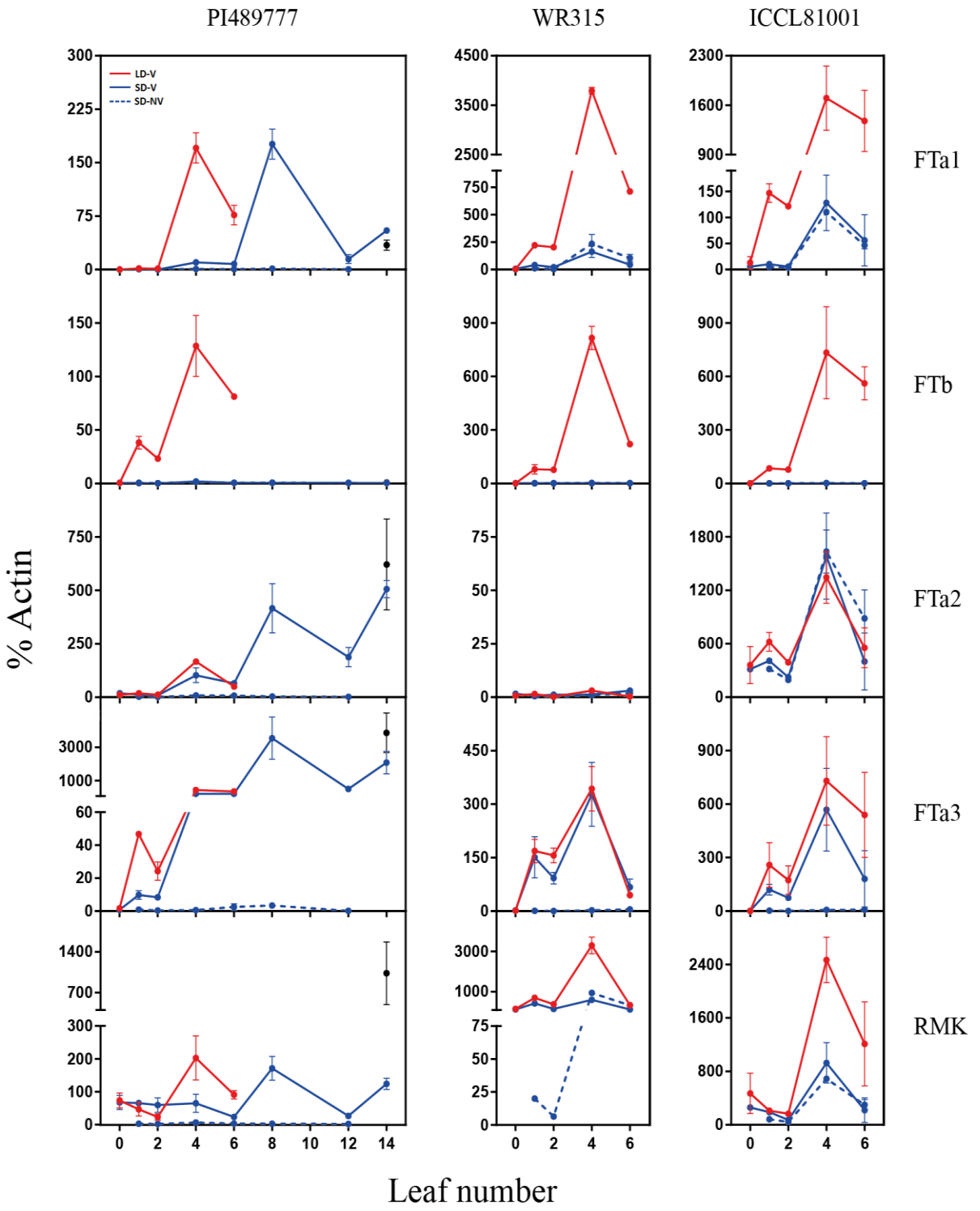


Figure 7.6 Expression of *FT* genes in leaf tissue of chickpea accessions PI489777 (wild), WR315 and ICCL81001. The unvernalized control treatment is represented by a blue dashed line, and vernalization treatments by solid lines, either SD (blue) or LD (red). Expression of the genes on revernalized leaf is indicated by black dot at leaf 14 only in PI489777 accession. Numbers on the horizontal axis represent the number of the leaf in which expression was determined, except for 0, which represents harvest of apical bud tissue one day after transfer of plants from cold to warm conditions. This sample is therefore absent in plots of NV data. Values on the Y axis represent expression levels relative to an *ACTIN* reference gene.

7.3.2 Vernalization response in lentil

7.3.2.1 Lentil accession ILL-2601 is vernalization-insensitive

The lentil accession ILL2601 has been catalogued amongst the earliest to flower in the lentil germplasm, and recent unpublished analysis of a cross between ILL2601 and the mid-late flowering accession ILL5588 has indicated the presence of a major locus conferring dominantly inherited early flowering that is located in the region of the *FTa1-a2* cluster on lentil linkage group 6 (Rajandran 2016). Differences in flowering were attributed to a ~10kb deletion in the *FTa1-FTa2* intergenic region in ILL2601 (Figure 7.7), and association analysis of the deletion with earliness in a worldwide collection of lentil accessions seems to corroborate this finding (Rajandran 2016). The similarity of this locus to the chickpea locus under investigation in this thesis suggested that the *FTa1-2* cluster might be regulated similarly in both species, and that the early variants might have a similar molecular basis. It was therefore of interest to examine how the *FT* genes might be regulated by vernalization and photoperiod in lentil and to investigate the effect of the deletion on this regulation.

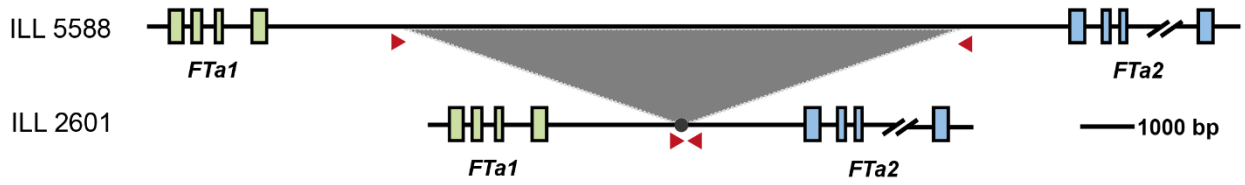


Figure 7.7 Representation of the *FTa1-FTa2* region in lentil accessions ILL5588 and ILL2601. Grey area represents the ~10kb deletion found in the intergenic region of ILL2601. Adapted from (Rajandran 2016).

As in the case of chickpea, suitable near-isogenic material was not available to analyse the effects of this locus, so the experiment was necessarily limited to a comparison of the parental lines. Vernalized and unvernalized plants from both accessions were grown under either LD or SD and assessed for flowering-related traits DTF (Days to Flower) and NFDF (Node of flower development).

In ILL2601, there was no significant effect of vernalization treatment on either DTF or NTF in either of in any of the two photoperiods tested, characterising this line as vernalization-insensitive (Table 7.5, Fig 7.8). In fact, under LD, vernalized plants actually flowered slightly later than unvernalized plants (2.4 days, $P<0.05$). In the late-flowering parent ILL5588, vernalization caused earlier flowering under both photoperiods. Under SD, unvernalized plants remained non-flowering after 123 days, whereas vernalized plants flowered after 71.8

days. Under LD plants flowered earlier and showed a small (2.3 days) but significant (P<0.05) promotion of flowering by vernalization. This interaction, where photoperiod influences the size of the vernalization response is similar to that seen in chickpea and suggests that vernalization and photoperiod may also share common molecular targets in lentil.

Table 7.5 Mean (μ) and standard deviation values obtained for days from emergency to first flower (DTF) and node of the first open flower (NFD) in lentil lines ILL2601 and ILL5588 grown in 4 different conditions.

			DTF		NFD	
Accession	**Treatment**	**N**	**Mean**	**Std. Dev**	**Mean**	**Std. Dev**
ILL2601	**NV-LD**	12	35.5	2.5	8.2	0.9
	V-LD	19	37.9	3.3	8.0	1.0
	NV-SD	9	65.7	5.7	10.8	0.8
	V-SD	10	62.8	8.6	10.6	1.3
ILL5588	**NV-LD**	13	41.4	2.8	13.6	1.2
	V-LD	21	39.1	2.5	12	1.5
	NV-SD	15	-[a]	-[a]	-[a]	-[a]
	V-SD	23	71.8	4.8	13	1.4

N= Number of plants; LD = Long days; SD = Short days; V = Vernalized plants; NV= Non-vernalized plants
-[a] Plants were unable to flower/set pods under this conditions during the scoring period.

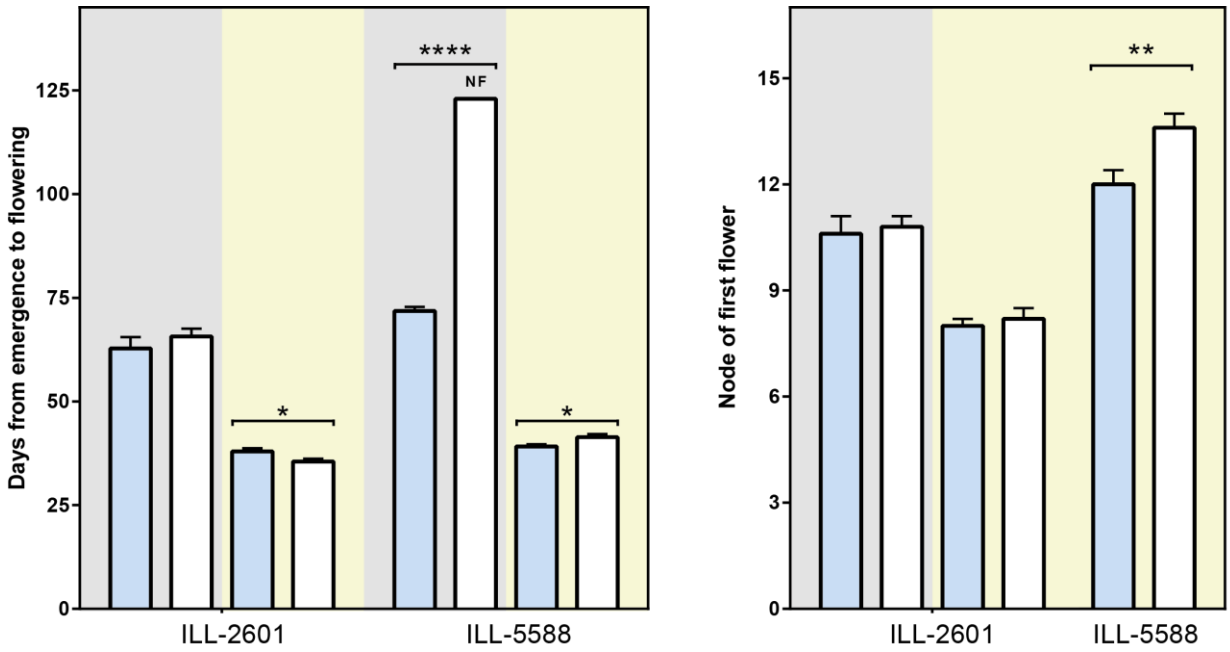


Figure 7.8 Phenotypic response of lentil accessions ILL2601 and ILL5588 to different combinations of photoperiod and vernalization treatment. Blue and white bars represent vernalized and unvernalized plants, respectively. Background colour represents photoperiod; short days are indicated in grey and long days in yellow. Unvernalized ILL5588 plants grown in SD were unable to flower (NF) but scored with a value of 123 days coinciding with the end of the phenotyping period. Consequently, their flowering node is not present in the corresponding graph.

7.3.2.2 Effect of vernalization on expression of lentil FT genes

Similar to the experiment described earlier in chickpea, the effect of expression of vernalization and photoperiod on lentil *FT* homologues was also examined in the same two lentil accessions. However, the experimental design was somewhat different, as illustrated in Fig 7.9, with vernalized and control plants grown under both LD and SD. In addition, in order to detect a possible induction *FT* expression during the cold treatment period, the entire apical bud (apical meristem and surrounding undeveloped leaves) was harvested in seedlings of both accessions at several timepoints during the 3-week vernalization treatment. Once plants were transferred to warm condition under SD or LD, leaves and shoot apex were independently collected from both vernalized and non-vernalized plants, one and two weeks after transfer. The material harvested for plant apex samples consisted of a combination of the apical meristem and the immature leaf tissue surrounding it. When harvesting between genotypes and across timepoints it was difficult to maintain consistency in the nature of the samples harvested, and it is possible that this may have introduced variability in the results and prevented clear conclusions.

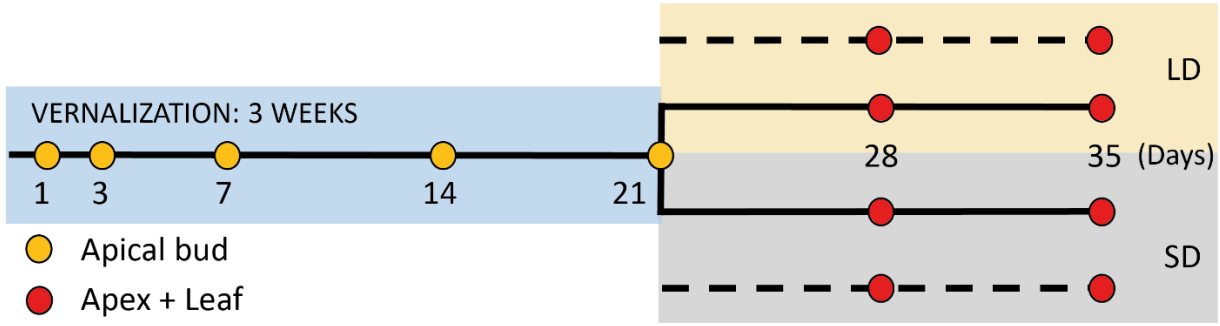


Figure 7.9 Graphical representation of the experimental design showing harvest points (filled circles) and tissues collected at each point. Numbers indicate days since transfer to cold room for vernalization. Continuous lines represent vernalized plants while dashed lines non-vernalized plants. SD = Short day; LD = Long day.

Figure 7.10 shows that during the cold treatment, the expression levels of all *FT* genes remained low and only the *LcFTa1-a2* intergenic region (*LcRMK*) seems to be induced above background by vernalization. After transfer to warm conditions, apical expression of *LcFTc* in the shoot apex was higher in vernalized than non-vernalised plants under SD. This difference in expression was observed in both accessions and was absent in long days.

In comparison, the leaf samples harvested after transfer to warm conditions showed better results. Overall, observations in lentil are in agreement with the picture described in chickpea;

LcFTb2 expression is restricted to leaf tissue and in both accessions is induced only by LD, while *LcFTa3* expression is induced by vernalization treatment after transfer to warm conditions, although some expression could be found also in unvernalized ILL5588 plants grown under LD. Also consistent with results obtained in cultivated chickpea, expression of *LcRMK* in the leaf appears responsive to both vernalization and long days.

No differential expression was found in *LcFTa2* or *LcFTc* in response to any factor. However, levels and expression pattern of *LcFTa2* showed some differences between the two accessions. Expression was substantially higher in ILL2601 than in ILL5588 in both LD and SD, and remained constant at both time points, whereas in ILL5588 expression was absent during the first week but rapidly increased in the second.

LcFTa1 expression levels in leaves of ILL2601 were in all cases much higher than comparable material from ILL5588, suggesting that this gene is substantially de-repressed in ILL2601 leaves. Also, ILL2601 levels were overall much higher in LD than in SD, indicating a strong regulation by photoperiod. In ILL5588, maximum expression is seen in response to the combination of vernalization and long photoperiod, suggesting that the cold treatment makes *FTa1* expression more responsive to the photoperiod stimulus.

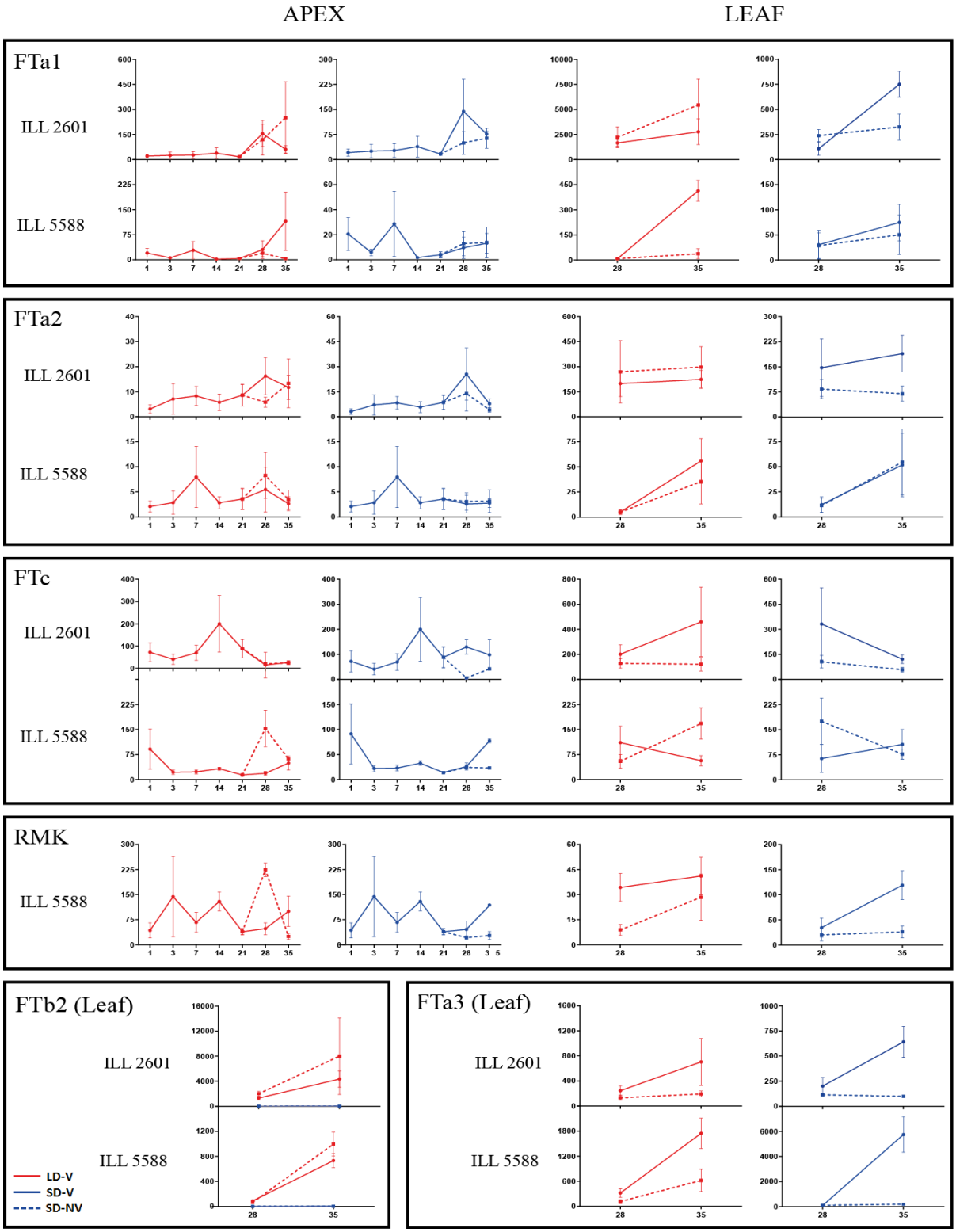


Figure 7.10 Expression of *FT* genes in lentil accessions ILL2601 and ILL5588. Results from vernalized plants are represented by continuous line and unvernalized plants by dashed lines. LD and SD conditions are indicated by red and blue lines, respectively. Numbers on the horizontal axis represent number of days from germination (1 to 21 correspond to vernalization treatment) and Y axis represent expression levels relative to an *ACTIN* reference gene. *LcRMK* is deleted in ILL2601 and its expression was therefore not tested.

7.4 Discussion

Vernalization response of cultivated chickpea

The vernalization response of cultivated chickpea is a controversial question. The first studies addressing this question concluded that a period of cold during the early stages of development promoted flowering to some extent, with the response becoming stronger as the photoperiod shortened (Nanda and Chinoy 1960; Pal and Murty 1941; Angus and Moncur 1980; Saxena and Siddique 1980). This conclusion was later challenged and the vernalization response proposed to be an artefact of experimental design and method of data analysis. Since that time chickpea has been considered a vernalization-insensitive species (Abbo et al. 2003a; Berger et al. 2012; Summerfield et al. 1989).

Our results suggest that all cultivated chickpea germplasm might retain the molecular pathways needed for response to vernalization. Under the short photoperiod used in this study, all cultivated lines show a clear and significant advance in flowering time after vernalization treatment. In this situation, vernalization can be as effective as photoperiod in floral induction. This substitution effect under extremely short photoperiods has been previously described (Angus and Moncur 1980). However, under long photoperiod, a bimodal response was found among cultivated accessions; some lines show a vernalization-insensitive phenotype, while other display a strong response that can be comparable to that observed in the wild accessions. This differential behaviour to vernalization between early and late lines is similar to that described in the literature, and suggests that that both vernalization and photoperiod pathway share common elements in their signalling pathway. Sharma and Upadhyaya (2015) found that vernalization accelerates the rate of progress toward flowering in medium and late-maturing chickpea varieties, while it had no or minimal effect in early-maturing lines. The use of early lines combined with non-restrictive conditions during the growing season could explain why other studies failed to detect the response, since it appears that in most of the cultivated germplasm the response may only become apparent under an extremely short photoperiod, such as the 8 h day used in this study.

Besides its effect on flowering time, vernalization was also found to consistently shorten the time between the opening of the first flower and the formation of the first pod in *C. reticulatum*. By contrast, this response was lost in *C. arietinum* accessions analysed; the advancement in podding time observed in the cultivated species was due to the early flowering of vernalized plants. This implies that flowering and maturity might have different

regulatory pathways and elements, some of which are altered in the cultivated germplasm making them irresponsive to vernalization.

Another effect of the vernalization treatment observed in this study is that reduces the number of aborted flowers. The presence of these abortive buds has been recently related with a low *FT* expression (Ridge et al. 2017), an idea consistent with the observations made in this study, as the lowest level of bud abortion was found in plants subjected to both long photoperiod and vernalization, conditions which also result in maximum expression of *FTa1*.

These findings could have profound implications for future breeding programs; several studies show that autumn-sown chickpea productivity is higher compared to that obtained when grown as a summer crop (Zaiter and Barakat 1995; Singh et al. 1997; Pinhasi van-Oss et al. 2016). It is therefore reasonable to assume that chickpea farming might return to the more productive winter cropping, once resistant cultivars can be developed. This return to autumn-sown chickpea would clearly require a re-introduction of winter adaptive traits lost by selective pressure during domestication of chickpea, including vernalization responsiveness (Berger et al. 2005). This introgression will necessarily need to be done from wild *Cicer* species, likely those in the primary (*C. reticulatum*) or secondary (*C. echinospermum*) gene pools (Abbo et al. 2002; Abbo et al. 2003a). However, the transfer of specific desirable traits from wild species to elite germplasm is a complex process due to the extensive genetic differences between them, and may be further complicated by the presence of tightly linked undesirable genes/traits in a phenomenon commonly referred to as "linkage drag". However, if a functional vernalization response within cultivated chickpea can be demonstrated and its genetic basis understood, this may provide an alternative source of responsive alleles in an approach that would be cheaper and simpler than the use of interspecific crosses.

Another aspect with agronomical significance is the possible independence between molecular regulation of flowering and maturity. This concept was already discussed in chapter 4, where two closely linked but distinct QTL regulating flowering time and maturity were described. The results obtained here add more evidence supporting this idea. Although the nature of these pathways remains unknown, the identification of elements exclusively involved in pod initiation could be of great importance, as early maturity alleles could be used together with early flowering alleles in the creation of short-cycle cultivars by gene

pyramiding. Such early cultivars would be invaluable in the short season environments where most of the chickpea is currently grown (Kumar and Abbo 2001).

Potential roles for FT genes in legume vernalization response

Molecular control of vernalization in plants is best understood in Arabidopsis and cereals. In both systems the mechanism of action at the most general level is similar; an epigenetic repression of flowering is overcome by a cold treatment, and this permits the promotion of flowering in response to photoperiod. This mechanism ensures that the reproductive development and seed production take place after winter has passed, and at a favourable time of year for plant growth and pollinator abundance (Greenup et al. 2009; Ream et al. 2012).

Since *FT* genes are well conserved in flowering plants as flowering activators, de-repression and induction of *FT* homologs is a common outcome of the vernalization pathway in different species: In both Arabidopsis and cereals, the silencing of *FLC* and *VRN2* genes enables *FT* induction by photoperiod (Dennis and Peacock 2009; Hemming et al. 2008). In sugar beet (*Beta vulgaris*), both floral repressor and activator roles are played by a pair of *FT* homologs (Pin et al. 2010). Finally, molecular studies in *Medicago truncatula* (Laurie et al. 2011; Jaudal et al. 2013) and *Lupinus angustifolius* (Nelson et al. 2017) all point to the conclusion that *FT* genes are involved in the vernalization response of legumes.

In the present study, *C. reticulatum*, *C. arietinum* and *L. culinaris* accessions were used to investigate the regulation of *FT* genes by vernalization and the extent to which this might be conserved in these two legumes. In the case of wild chickpea all three genes belonging to the *FTa* clade (*CrFTa1*, *CrFTa2* and *CrFTa3)* show induction in response to vernalization while *CrFTb* is induced exclusively by long photoperiod. Results obtained in lentil accession ILL5588 are consistent with those described in chickpea; *LcFTb2* is the main target of photoperiod and *LcFTa* genes (except *LcFTa2*) were responsive to vernalization. In Medicago, *MtFTa1* and *MtFTa2* also show elevated expression in response to vernalization (Laurie et al., 2011), although expression of *MtFTa3* has not yet been tested in this species. Among the three *FTa* genes, *FTa1* seems particularly relevant in both chickpea and lentil, as the only one that responds to both photoperiod and vernalization. The role of *CrFTa1* as integrator of both pathways is also a common feature with Medicago, as revealed by previous work in this model legume (Laurie et al. 2011).

One interesting and potentially important result concerns the expression of *CrFTa3* and *CrFTb*. These genes show specific induction in response to vernalization (*CrFTa3*) or LD (*CrFTb*), respectively, and in both cases this induction appears to occur somewhat earlier than induction of *CrFTa1* and *CrFTa2*. This suggests that *CrFTa3* and *CrFTb* may be specific targets of vernalization and photoperiod regulation, respectively, and may help mediate specific flowering responses to these factors.

Finally, the chickpea *FTa1-FTa2* intergenic region (*CrRMK*) shows an expression pattern mirroring that of *CrFTa1*, with the notable difference that levels of *CrRMK* are high even at the earliest developmental stage in vernalized plants (1 day after vernalization treatment). Interestingly, leaves developed under the cold show a significant increase in the expression of *CrRMK*. In lentil, *LcRMK* is the only tested gene that seems to be expressed by cold in the apical bud of plants harvested during vernalization treatment, with a peak of maximum expression after 2 weeks. These observations suggest that expression of this region can be directly induced by cold in both species, in contrast with *FTa1* itself which may be released from repression by cold, but is expressed only on return to warm conditions. The cold-induced expression of *RMK* is maintained after transference to warm; in both chickpea and lentil, foliar *RMK* expression was higher in vernalized plants and it also seems responsive to long photoperiods.

Flowering phenotype alterations in accessions from different species showing polymorphisms in this region further support its importance in the control of vernalization; the lentil accession ILL2601, bearing a ~10 kb deletion that completely eliminates *LcRMK*, shows a vernalization-insensitive but still photoperiod-responsive phenotype. This is a similar flowering behaviour to that described for Medicago "*spring"* mutants (Jaudal et al. 2013), caused by retrotransposon insertions in the 3' region of *MtFTa1*. Interestingly, the four chickpea lines showing a weaker vernalization response (ICCV2, WR315, ICCL81001 and ICC4958) present either a deletion or an insertion in this region. Taken together, these results suggest a possible involvement of the region between *FTa1* and *FTa2* in the vernalization process that is conserved across all galegoid clade of legumes.

Molecular basis for flowering time QTL in FTa1-2 cluster region

In both chickpea and lentil, major flowering time QTLs are located over the *FTa1-a2-c* cluster, suggesting these genes as candidates. In chickpea, the intrinsic upregulation of these genes in domesticated accessions was described in chapter 4, and together these results

suggest that one or more genes in this cluster may be responsible for the early flowering of *C. arietinum* compared to *C. reticulatum*.

Here, we present evidence that *FTa1* is a key flowering gene and a common integrator for vernalization and photoperiod in wild chickpea. The early, cultivated lines used in this study (WR and ICCL81001) show no alteration in the regulatory patterns of those *FTs* located in other genomic regions (*FTb* and *FTa3*), whereas the expression profile of *FTa1* shows early induction after the second leaf independent of the photoperiod or vernalization treatment. The intensity of this induction, however, is enhanced by long photoperiods, as *FTa1* levels are considerably higher in LD than in SD, which can explain the differences in flowering time observed between the two photoperiods in these accessions. Overall, these differences represent a major contrast in the behaviour of this region in domesticated compared to wild chickpea, where no expression of any of the genes was ever observed in unvernalized plants.

The molecular basis for this *FTa1* upregulation are unknown, but our results suggest that is likely due to the partial loss of a vernalization requirement. Since *RMK* seems to play a role in the regulation of vernalization pathway, it is plausible that changes in this region could be involved in this upregulation. A simple model explaining these observations is proposed in figure 7.11; first, it assumes that transcription of *RMK* is coupled to that of *FTa1* but not *vice versa*, based on the identical regulation patterns of *FTa1* and *RMK*, except for the above mentioned direct cold-induced *RMK* (not seen in *FTa1*). Second, as discussed in the previous section, disruption of *RMK* in different species appears to be associated with by a reduction or loss of the vernalization response, and an upregulation of *FTa1* in the absence of vernalization. This suggest the presence of an *FTa1*-repressive element in the *FTa1-a2* intergenic region that is the target of either epigenetic marks or a mobile repressive element, like a transcription factor. Since the capacity for vernalization response needs to be reset in the next generation, it is likely that epigenetic regulation is common feature even, so the first option is more likely to be true (Yuan et al. 2016; Kim et al. 2009; Amasino 2004). In any case, the specific induction of *RMK* expression (by cold) or *FTa1* expression promoted by any other factor (e. g. photoperiod) could potentially override this repression.

This is the first model for the molecular regulation of vernalization in chickpea. To date, the only attempt to map loci associated with vernalization response in chickpea is that of (Samineni et al. 2016), who found a major QTL in LG3 controlling this trait. Not surprisingly,

the *FTa1-a2-c* cluster lies within the interval defined for this QTL, making it likely that it is the same one investigated in this thesis.

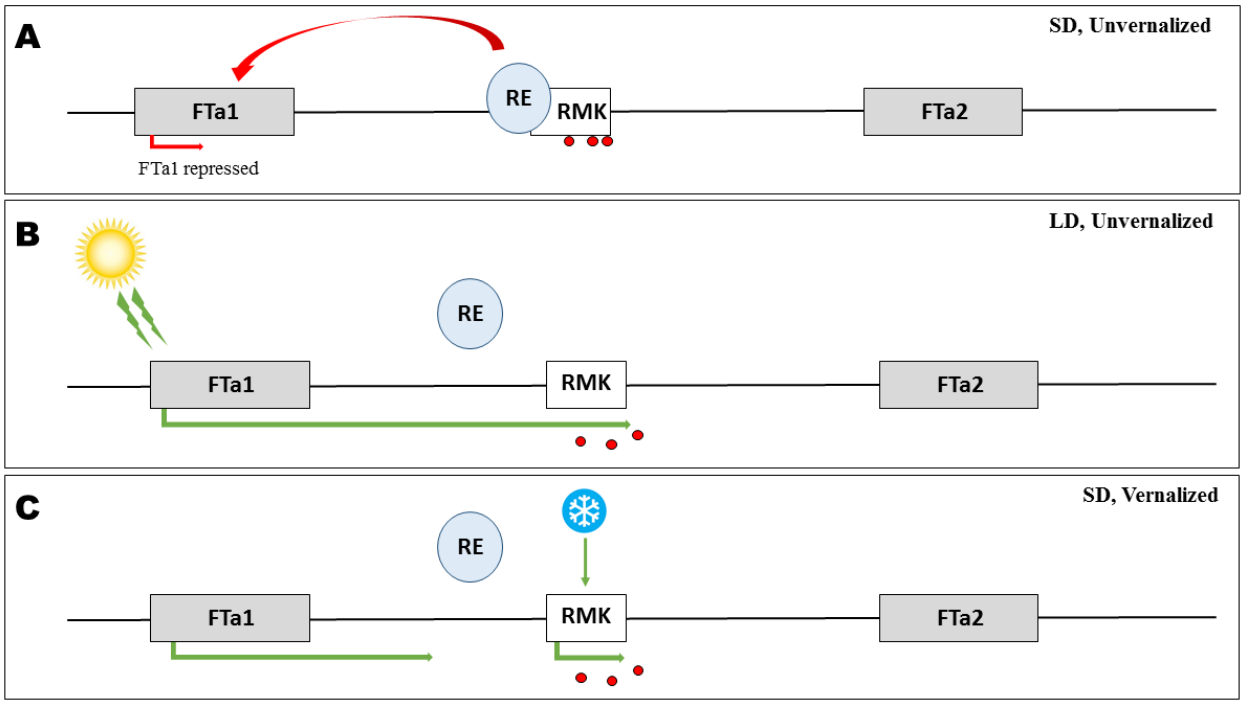


Figure 7.11 Model of interaction of *FTa1* and *RMK*. In absence of inductive factors (**A**), regulatory site(s) at *RMK* are occupied by either repressive epigenetic marks (red dots) or repressor elements (RE) acting in trans. Induction of *FTa1* expression by other factors, such as photoperiod (**B**) eliminates the repressive elements at *RMK* due to the coupled expression of both genes. Direct induction of *RMK* expression by vernalization (**C**) may have a similar effect.

In lentil, the upregulation of *LcFTa1* and *LcFTa2* in the line ILL2601 was proposed to be the main cause of its early flowering phenotype (Rajandran 2016). It was also hypothesised that the deletion of around 10 kb in the intergenic *FTa1*/*FTa2* region could be the cause of this upregulation. Observations in this study are consistent with this conclusion, as the expression levels of these two genes were high in ILL2601 leaves independent of the conditions in which plants were grown. Results also show that this accession is vernalization-insensitive, so it is likely that earliness in this line is due to the loss of vernalization requirement. Finally, the model above offers a likely explanation to how this deletion can lead to the upregulation of neighbour *FT* genes.

Conclusion

Results presented in this chapter for both chickpea and lentil show that *FTa3* and the *FTa1*/*a2* intergenic region (*RMK*) are regulated by vernalization and likely to participate in the vernalization response in both species. *RMK* seems to be induced directly by cold and might

be indispensable for a proper vernalization response, while *FTa3* requires transfer to warmer conditions for its expression. *FTa1* is the only gene in both species whose expression levels respond to both photoperiod and vernalization. The role of *FTa1* as integrator of different cues seems to be shared, at least among legumes of the galegoid clade (Laurie et al. 2011; Hecht et al. 2011; Jaudal et al. 2013), suggesting a conserved molecular mechanism for vernalization in all cool-season legumes.

Chapter 8. General discussion

8.1 Summary of main findings

A short growing season is currently one of the major constraints for chickpea productivity globally. Consequently, a major goal of chickpea breeding is the development of short-duration varieties to fit these environments. Flowering time is a critical contributor to overall crop duration, and thus flowering time (earliness) is a key target in many breeding programs. Traditional breeding approaches have had significant success in developing early-maturing varieties, but there is little understanding of the specific genetic and physiological mechanisms by which this earliness is achieved. Newer technologies based on molecular biology, genomics and bioinformatics offer enormous potential both in identifying the molecular basis for early phenology and in the design and efficient development of new varieties tailored to specific locations and situations. In the case of complex traits such as flowering time, a commonly adopted first step towards these aims is the use of association analysis to identify chromosomal regions associated with the regulation of a particular trait. The regions thus defined can later be the subject of more detailed studies to uncover the identity of the genes governing the targeted trait. In chickpea, numerous association studies in the last decades have addressed flowering time (summarized in Table 1.2), and have identified several regions important for floral induction, among which the central region of chromosome 3 is particularly prominent (Cobos et al. 2009; Mallikarjuna et al. 2017; Hossain et al. 2010; Rehman et al. 2011; Aryamanesh et al. 2010). Moreover, comparison of flowering QTLs in other temperate legumes has revealed that this region is syntenic with regions important for flowering control in several other species, suggesting the presence of a major flowering locus conserved across this group of legumes (Weller and Ortega 2015). In the same review, we proposed that in view of their widely conserved role as floral activators, a cluster of *FTa* and *FTc* genes in this region represent the most likely candidates for these QTLs.

The primary objective of this thesis was to investigate this possibility. First, we used the newly-available chickpea genome sequence of CDC Frontier to a) characterize the chickpea family of PEBP genes, focusing on the *FT* genes (Fig 3.2, Fig 3.3), b) study the conservation in chickpea of 234 Arabidopsis flowering genes and c) map those markers delimiting flowering QTLs reported in the eight chromosomes. Combining all this information, we built a chickpea genome map showing the co-location of 241 potential candidate genes with the

published flowering QTLs (Fig 3.5). Using this approach, we delimited the above-mentioned region of chickpea chromosome 3 between markers TA6 and TA64 and confirmed the position of the *FTa-c* genes within this interval.

In chapter 4, we reported twelve QTL for flowering time within the TA6-TA64 interval using four chickpea recombinant inbred populations grown in different conditions and seasons (Table 4.5). The co-location of some of these QTL identified two distinct relevant regions; the first, between markers LOB189 and PRT6, was found to be relevant only in the intraspecific populations (Fig 4.6 C, D; Fig 4.7). The second region is defined by markers CDF2d and SUVH4, and harboured seven QTL in both inter and intraspecific populations (Fig 4.6 A, B; Fig 4.7). The importance of this region in the control of flowering appears to be different between species. Whereas it controls most of the phenotypic difference in flowering time between *C. arietinum* and *C. reticulatum* (Fig 4.6**)**, its effect was weaker in the intraspecific cross and appears to be significantly influenced by environmental conditions. Using fine mapping, we narrowed this interval to a 0.8 Mb region containing 59 annotated genes, including the *FTa-c* cluster (Fig 4.4). Furthermore, a marker for the *FTa1* gene was the most strongly associated with the phenotype in this region among all the markers tested, consistent with our initial hypothesis.

Results from QTL analysis were further supported by differential expression of *FT* genes in the cluster, which were upregulated in the early parents of the crosses when compared to late ones (Fig 4.8 and Fig 4.9). This result is consistent with an expected gain-of-function mutation that would explain the early flowering of *C. arietinum*. The lack of *FTa2* in the accession WR315 leaves *FTa1* and *FTc* as the more likely candidates. Furthermore, the fact that the QTL effect in the CRIL2 population was strong under SD but could not be detected under LD is consistent with the known physiological roles of the *FT* genes as integrators of different flowering signals including photoperiod. In chickpea, temperature and photoperiod are the main determinants of flowering time (Roberts et al. 1985). Therefore, under inductive photoperiods, any differences in *FT* alleles are not relevant as these genes are already upregulated by these conditions. Only in absence of other promoter signals, does the intrinsic, endogenous upregulation caused by differences in the sequence of the genes become more evident. Based on these findings and the fact that QTL reported to control flowering in other interspecific populations coincide with this region, we proposed that this region contains a domestication locus that is the predominant factor responsible for the altered phenology of *C. arietinum* compared to its wild ancestor.

In addition to QTL in the two central regions of chromosome 3, a third QTL was detected in the CRIL2 population, in a distinct position very close to the bottom of LG3 (Fig 4.6 A), which may represent a novel locus. Finally, two other QTL were found in LG4 (Fig 4.5), associated with markers STMS11 and GAA47 in a region of the chromosome previously reported to be associated with the control of flowering (Mallikarjuna et al. 2017; Cobos et al. 2007; Pushpavalli et al. 2015; Daba et al. 2016a; Varshney et al. 2014b). All these regions could have important implications for manipulation of phenology, by allowing the directed pyramiding of earliness genes to create genotypes better adapted to short-season environments.

Moderate to very strong correlation was found between flowering, growth habit and branching phenotype in the interspecific population CRIL2, in particular between flowering and growth habit (Fig 5.3). QTL analysis mapped the genetic control of growth habit and branching to the same region as the flowering time QTL (Fig 5.4), with no clear evidence obtained for recombination between flowering time and growth habit, suggesting that these might be pleiotropic effects of a single locus. *Hg* is a previously-described major locus controlling growth habit between wild and cultivated chickpea, and has been mapped in previous reports to a region of chromosome 3 containing the *FTa-c* genes (Muehlbauer and Singh 1987; Kazan et al. 1993; Winter et al. 2000). Collectively these observations suggest that the *FTa-c* cluster are also strong candidates for *Hg*. This conclusion is also supported by observations that that *C. reticulatum* maintains apical dominance when grown under LD or after being vernalized, whereas this feature is lost in non-inductive flowering conditions, leading to a prostrate habit and profuse branch outgrowth **(**Fig 5.1). Similar responses have been documented by Abbo et al. (2002). Since both vernalization and photoperiod pathways appear to act through *FT* genes in chickpea to promote flowering (Chapter 7), these observations could be explained by an upregulation of *FT* genes and thus variation in the *FT*/*TFL1* ratio. This ratio has been proposed to regulate different developmental processes in many species (Shalit et al. 2009; Lifschitz 2014), including branching pattern in rice (Tamaki et al. 2007), rose (Randoux et al. 2014) and tomato (Lifschitz 2008). The opposite expression profiles obtained for the *FT* and *TFL1* chickpea homologs (Fig 4.8 to 4.10**)** are consistent with this hypothesis. Further support from results in Medicago, where *FTa1* mutants showing elevated expression of this gene shows longer primary axis and fewer branches, and *vice versa* (Jaudal et al. 2013; Laurie et al. 2011).

The elevated expression of *FT* cluster genes in lines carrying the domesticated allele at the major LG3 QTL suggested the possible presence of mutations in or around these genes that might de-repress their expression. Hundreds of polymorphisms were found between the *FTa1*, *FTa2* and *FTc* alleles of wild and cultivated chickpea, and the promoter region of *FTa1* was particularly variable (Fig 6.4). Although it is impossible to determine which of these might be causal for the expression and flowering time effects without further experimentation, many of these mutations could potentially eliminate repressor sites leading to upregulation of the genes, as reported in *FT* genes in other species. Diversity in the *FT* cluster sequence within the domesticated germplasm was much lower, as expected in a crop species that suffered four bottleneck events during its evolution (Abbo et al. 2003a). However, significant insertions and deletions were nevertheless identified and found to show some degree of association with flowering time differences (Fig 6.8, Fig 6.9). Although these preliminary results need to be confirmed in larger and better-characterized populations, the sequences obtained in this study are invaluable to the development of markers targeting these genes.

The results obtained in chapter 7 demonstrate that *C. arietinum* is a vernalization sensitive species, supporting the results obtained recently by (Sharma and Upadhyaya 2015; Pinhasi van-Oss et al. 2016), and suggest that the general perception that chickpea as a vernalization unresponsive species should be revised. Two distinct types of vernalization response were found among the 94 accessions analysed; in one group, flowering promotion by vernalization is apparent under both LD and SD, and is comparable in magnitude to the response of *C. reticulatum*, whereas in the second group, characterized by earlier flowering overall, the response to vernalization is only seen under SD and absent in LD (Fig 7.2). This indicates that both cues share common components in their signalling pathways. Vernalization also reduced the time needed for the appearance of the first pod (Fig 7.3), although this was a consequence of the early flowering in the case of cultivated accessions, as the time from opening of the first flower until the appearance of of the first pod was similar (Fig 7.4). The expression analysis points to *FTa1* as the likely integrator of both photoperiod and vernalization (Fig 7.6). Evidence was also obtained for the likely involvement of the *FTa1-FTa2* intergenic region in the regulation of vernalization, a feature that seems conserved in temperate legumes as disruption of this region in Medicago can also reduce or eliminate vernalization response (Jaudal et al. 2013), and we obtained similar results in the lentil line ILL2601 (Fig 7.8). A long non-coding RNA (lncRNA) is present in the region, and we present evidences that its expression could be directly upregulated by cold (Fig 7.6, Fig 7.10),

suggesting a possible role in vernalization response of legumes. Collectively these results provide the first insights into the possible molecular basis for vernalization in chickpea. They are also complemented by the recent study of Samineni et al. (2016), who recently examined the genetic control of vernalization response in chickpea and found that more than half of the phenotypic variation was controlled by a region in chromosome 3 that coincides with the location of the *FT* cluster investigated here.

8.2 Flowering model in chickpea

Based on the results obtained in different chapters of this thesis and previous works on the molecular basis of flowering in chickpea and other legumes, I have developed a model to illustrate the flowering regulation on chickpea.

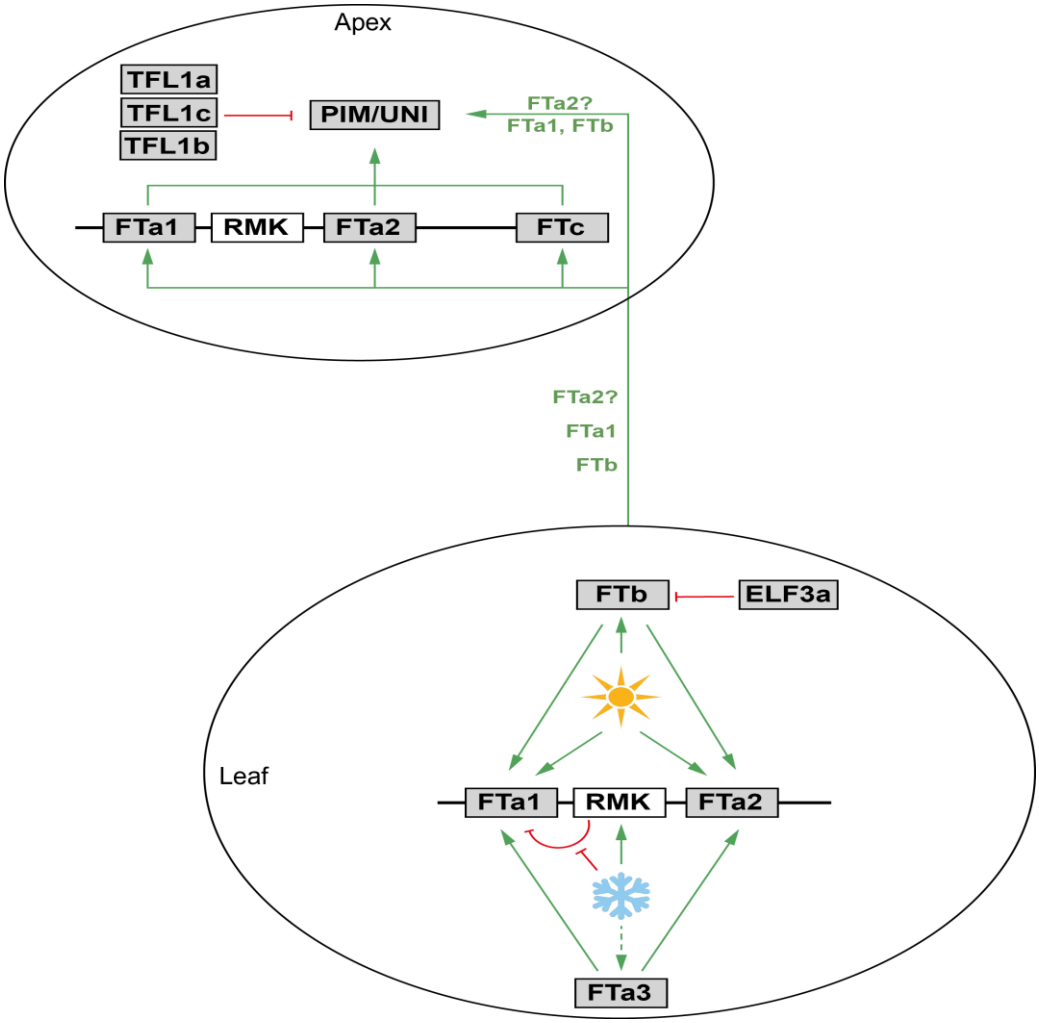


Figure 8.1 Model of *FT*-dependant floral induction in chickpea. Green lines represent induction and red line repression.

The observed expression patterns of meristem identity genes such as *PIM* and *UNI*, the chickpea orthologs of Arabidopsis meristem identity genes *AP1* and *LFY*, are consistent with a conserved role downstream of *FT* genes (Fig 4.8 and Fig 4.10),. The fact that this expression occurs prior to appearance of flower buds makes them good indicators of floral commitment, similar to the well-studied pea system (Hecht et al. 2011; Ridge et al. 2017). Collectively, the results obtained in different chapters of this thesis suggest that all five chickpea *FT* genes could potentially participate in the floral induction in chickpea. Their distinct profiles of expression in different tissues and in response to environmental cues indicates a probable divergence in their physiological roles, again consistent with other temperate legumes (REFS). The single chickpea *FTb* gene is exclusively expressed in leaves and is only induced by long photoperiod in both wild and cultivated chickpea species (Fig 4.9, Fig 7.6). The only clear evidence for its regulation comes from the recent study of Ridge et al. (2017) who showed that its specific repression under SD requires the circadian clock gene *ELF3a*. Although *FTb* genes have been proposed as important and specific targets of the photoperiod pathway in temperate legumes, there is so far no direct evidence demonstrating this role.

In contrast, *FTa1* has been analysed functionally and has major effects on flowering time in both pea and Medicago, and the results here suggest that chickpea *FTa1* may also have an important role. In Medicago, *FTa1* expression integrates signals from both photoperiod and vernalization (Hecht et al. 2011; Laurie et al. 2011), and very similar results were obtained for chickpea (Fig 4.9, Fig 7.6). Interestingly, in Medicago, loss of *FTa1* eliminates the response to vernalization but not photoperiod, indicating that it may be critical for vernalization response but that response to photoperiod may be mediated by additional *FT* genes, most likely FTb genes. In wild chickpea, the timing of *FTa1* induction follows that of *FTb* in LD and with that of *FTa3* under SD when vernalized. The earlier expression of these genes suggest they could be primary *FT* targets of light (*FTb*) and cold (*FTa3*). One possibility is that they might then be involved in the induction of *FTa1*, in a cross regulation mechanism between *FT* genes that has been discussed by Hecht et al. (2011) in relation to the pea system. However, it is also possible that vernalization and photoperiod response mechanisms may be integrated directly at *FTa1* could also be possible, and further research involving different *FT* mutants and mutant combinations would be necessary to elucidate these interactions.

The regulation patterns of *FTa2* perfectly correlated with *PIM*/*UNI* induction in both apex and leaf tissues of *C. reticulatum* (Fig 4.8 Fig 4.9 and Fig 7.6). However, in *C. arietinum* the expression of this gene is totally altered. Deletions that completely eliminate the gene are present in ~40% of the total cultivated germplasm, while in the remaining lines, a retrotransposon insertion is associated with consititutive upregulation of *FTa2*. Despite these major polymorphisms no major alteration of flowering time was consistently observed in any case, suggesting that this gene, if it does have any function, may be redundant and not essential for floral induction. However, its lack of influence on flowering time cannot be conclusively demonstrated without further experiments, as both deletions were associated with a slight delay in flowering under LD, and *type2* shows potential earliness in SD (Fig 6.10). Since these deletions also affect both *FTa1*-*FTa2* and *FTa2*-*FTc* intergenic regions, it is possible that the observed effects might instead be due to loss of regulatory sequences within these regions. In this respect, recent evidence suggests that disruption of the genomic region between *FTa1* and *FTa2* may influence flowering time through an effect on the adjacent *FTa1* gene. In Medicago, retroelement insertions in this region confer early, vernalization unresponsive flowering and elevated *FTa1* expression (Jaudal et al. 2013). In this study, the lentil accession ILL2601, which carries an ~10kb deletion in this region (Rajandran 2016), shows a similar flowering time phenotype and *FTa1* expression (Fig 7.8). The fact that in both cases the "mutant" plants are still able to respond to photoperiod, indicates that the *FTa1*-*a2* intergenic region may be involved exclusively in the vernalization response, in both species. Although the molecular basis of its action is still unclear; a long non-coding RNA (ncRNA, called *RMK* throughout this thesis) can be found in chickpea and lentil in this region (Rajandran 2016) and is annotated in recent Medicago genome builds. In chickpea, the expression profile of this ncRNA mirrors that of *FTa1* in both wild and cultivated species, so it seems likely that *RMK* transcription is largely coupled to that of *FTa1* (Fig 7.6). But unlike *FTa1*, it seems directly induced by cold, consistent with a possible role in the early steps of the vernalization response at this locus (Fig 7.6). One possible model of action could involve the presence of one or more sites in the *FTa1*/*a2* intergenic region that are targeted by repressive factors such as DNA or histone modifications or transcription factor that downregulates *FTa1*. Induction of *RMK* expression by vernalization would eliminates such repression (Fig 7.11).

Another component in the genetic network governing vernalization response in chickpea could be *FTa3*, according to its regulation pattern. The expression of this gene was induced

after a period of cold and independent of photoperiod. Unlike *RMK*, it is not induced directly to cold, but it is inmediately upregulated after transference to warm conditions (Fig 7.6).

Finally, *TFL1* genes are floral repressors, preventing the apical meristem to floral meristem transition in a wide range of species, from arabidopsis (Bradley et al. 1997) to legumes such as pea (Foucher et al. 2003) and soybean (Tian et al. 2010). In chickpea, there are three *TFL1* homologs, namely *TFL1a*, *TFL1b* and *TFL1c*. The expression patterns of *TFL1a*, obtained by Ridge et al. (2017), are consistent with a repressive role in flowering, showing higher levels at early stages and decreasing with plant development. A similar expression profile was obtained for *TFL1b* and *TFL1c* in this study (Fig 4.10), and these results are in general consistent with those obatined in pea (Hecht et al. 2011), which indicates that the role of these genes could be conserved among legumes. The function of *TFL1c* may be particularly relevant, since in pea mutations in the gene also confer early flowering independent of environmental conditions (Foucher et al. 2003). Once again, specific functional studies will be required to confirm their role in chickpea.

8.3 Future directions

The model presented above represents an improvement in the understanding of the control of flowering in chickpea. However, it is essentially a broad framework to guide and interpret future studies, and incorporates several different hypotheses that will require detailed experimental evaluation. Some of the opportunities for future research are summarised below:

Further work aimed at confirming a role for the *FTa-c* cluster in flowering time and shoot architecture QTL should be a high priority, given its apparent major importance. One useful approach will be the development of near isogenic lines differing only in the cluster, which will allow a more accurate analysis of the correlation between *FT* regulation and flowering phenotype, free from the influence of other genomic regions. Another valuable but challenging strategy would involve the isolation of loss-of-function mutants for *FT* genes in the cluster, either by screening of chickpea natural diversity, or by targeted disruption using genomic editing techniques, such as the CRISPR/Cas9 system. These natural or artificial mutants would enable the specific roles of each gene to be definitively tested. In the particular case of *FTa2*, finding a knockout mutant (such as a nonsense mutation) that impairs its function without affecting surrounding regions would be ideal to evaluate the significance of the natural deletions identified in this study. If the relevance of this gene can be demonstrated, then the alternative splicing identified in Chapter 6 raises several questions

that would warrant further investigation, such as whether the various isoforms may be differentially regulated through development, in different tissues, or in response to different environmental conditions.

Chapter 4 also identified a second, novel region in LG3 that influences flowering time in the intraspecific populations. None of the flowering-related genes evaluated in chapter 3 was found within this 4Mb interval. More detailed evaluation of the 244 genes annotated in the region will now be required to select likely candidates, and narrowing the interval through design of new markers and/or increasing the population size will also be valuable. Phenotyping the existing population under controlled conditions may also help amplify the phenotypic effect and refine the estimate of the QTL interval.

One major hypothesis that emerges from this thesis is that that the *FT* cluster, and in particular the *FTa1* gene and its associated regulatory sequences, could be a locus that has played a significant role in the domestication of *C. arietinum*, through effects on flowering time and associated pleiotropic effects on shoot architecture traits. This hypothesis is supported not only by results in this thesis but also by multiple studies of interspecific populations in which major QTL appear in the same region: studies that provide some degree of "replication" to our observations. However, the number of such populations is small, and a first step towards the validation of this idea would be the development of more interspecific populations using wild and domesticated parental lines. In the case of domesticated parents, it would be most informative to use diverse landraces representing the early dispersal patterns of chickpea, as deduced from archaeological and genetic diversity studies, and also different climatic, altitudinal and latitudinal origins.

In chapter 6, we found promising associations between some polymorphisms and flowering time. However, the small size of the population used in the present study makes them inconclusive. A wider survey of these polymorphisms in the global chickpea germplasm backed by detailed phylogenetic information and more extensive phenotyping will be needed to accurately determine their potential significance.

Chapter 7 demonstrated that *FTa1* and several other *FT* genes are induced by vernalization treatment. However, this induction occurs only after transfer of plants to warm conditions, consistent with the general understanding of the vernalization response as an epigenetic phenomenon that requires cell division for its expression. It also indicates that *FTa1* induction may not be the primary regulatory step in response to cold but is rather likely to be

the key functional target. In contrast, the lncRNA in the *FTa1-FTa2* intergenic region is upregulated during the cold exposure period suggesting that it could have a role upstream of *FTa1* in mediating cold response and epigenetic modifications. Such a role might be analogous to the regulation of *FLC* by the *COOLAIR* lncRNA in Arabidopsis. Further investigation of the role and interaction of this lncRNA is likely to be challenging, but the first steps might be to better define its structure and attempt to specifically disrupt it in some manner. The availability of the chickpea genome and the high-quality sequence across the *FT* cluster generated in this thesis open the possibility of using chromatin immunoprecipitation (ChIP) approaches to characterize how vernalization and the LG3 QTL affect epigenetic histone modifications.

Chapter 7 also touched on an apparent differential effect of vernalization on the the number of days between the opening of the first flower and the appearance of the first pod. This is an important trait in the field, where maturity can be delayed and yield impaired by poor pod set. The delay in pod formation was significantly reduced in *C. reticulatum* in response to vernalization, whereas in the cultivated species this trait seems to be more dependant on ambient temperature. Whether this latter observation is true needs to be confirmed in controlled-temperature studies. In any case, genetic and molecular basis of the strong pod set in the wild species should be elucidated in order to understand its control and to enable its introduction to domesticated germplasm. This would likely have substantial benefit in the development of early-maturing cultivars to fit the short season enviroments, particularly those that can experience low temperatures during the pod development stage.

8.4 Conclusion

Overall, the work in this thesis has contributed to understanding of the mechanisms governing floral induction in chickpea, and provides new insight into the role of the *FT* genes in this process. It has identified several aspects of flowering genetics, physiology and molecular biology that will be valuable to explore further in the future, and has helped to formulate some new perspectives on diversity of flowering time in chickpea, and the evolution of flowering time adaptation in chickpea and other related temperate legumes. Ultimately this should assist with a number of applied goals in chickpea improvement, and will be particularly relevant in improving the efficiency of access to the much wider range of genetic variation in wild chickpea species. This contribution is particularly timely because it comes at a time when resources and tools for chickpea research are rapidly expanding and

making possible a whole new range of approaches and practical applications for understanding and harnessing genetic diversity.

References

Abbo S, Berger J, Turner NC (2003a) Evolution of cultivated chickpea: Four bottlenecks limit diversity and constrain adaptation. Funct Plant Biol 30 (10):1081-1087.

Abbo S, Lev-Yadun S, Galwey N (2002) Vernalization response of wild chickpea. New Phytol 154 (3):695-701.

Abbo S, Saranga Y, Peleg Z, Kerem Z, Lev-Yadun S, Gopher A (2009) Reconsidering domestication of legumes versus cereals in the ancient near east. The Quarterly review of biology 84 (1):29-50.

Abbo S, Shtienberg D, Lichtenzveig J, Lev-Yadun S, Gopher A (2003b) The Chickpea, summer cropping, and a new model for pulse domestication in the ancient Near East. Quarterly Review of Biology 78 (4):435-448.

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science 309 (5737):1052-1056.

Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F (2010) cis-Regulatory Elements and Chromatin State Coordinately Control Temporal and Spatial Expression of FLOWERING LOCUS T in Arabidopsis. The Plant Cell 22 (5):1425-1440.

Aguilar-Martínez JA, Poza-Carrión C, Cubas P (2007) Arabidopsis Branched1 acts as an integrator of branching signals within axillary buds. Plant Cell 19 (2):458-472.

Ahmad F, Slinkard AE (1992) Genetic relationships in the genus Cicer L. as revealed by polyacrylamide gel electrophoresis of seed storage proteins. Theor Appl Genet 84 (5-6):688-692.

Ahn JH, Miller D, Winter VJ, Banfield MJ, Jeong HL, So YY, Henz SR, Brady RL, Weigel D (2006) A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. Embo J 25 (3):605-614.

Alexandre CM, Hennig L (2008) FLC or not FLC: the other side of vernalization. Journal of Experimental Botany 59 (6):1127-1135.

Ali L, Azam S, Rubio J, Kudapa H, Madrid E, Varshney RK, Castro P, Chen W, Gil J, Millan T (2015) Detection of a new QTL/gene for growth habit in chickpea CaLG1 using wide and narrow crosses. Euphytica 204 (2):473-485.

Allchin FR (1967) Early Cultivated Plants in India and Pakistan. Duckworth,

Amasino R (2004) Vernalization, Competence, and the Epigenetic Memory of Winter. The Plant Cell 16 (10):2553-2559.

Amasino RM, Michaels SD (2010) The timing of flowering. Plant Physiology 154 (2):516-520.

Anbessa Y (2006) GENETIC ANALYSIS OF EARLINESS TRAITS IN CHICKPEA (CICER ARIETINUM L.). Thesis.

Anbessa Y, Warkentin T, Bueckert R, Vandenberg A (2007) Short internode, double podding and early flowering effects on maturity and other agronomic characters in chickpea. Field Crops Research 102 (1):43-50.

Anbessa Y, Warkentin T, Vandenberg A, Ball R (2006) Inheritance of time to flowering in chickpea in a short-season temperate environment. Journal of Heredity 97 (1):55-61.

Andrés F, Coupland G (2012) The genetic basis of flowering responses to seasonal cues. Nature Reviews Genetics 13 (9):627-639.

Angus JF, Moncur MW (1980) Photoperiod and vernalization effects on phasic development in chickpea. International Chickpea Newsletter (No.2):8-9.

Araki T, Komeda Y (1993) Analysis of the role of the late-flowering locus, GI, in the flowering of Arabidopsis thaliana. The Plant Journal 3 (2):231-239.

Aryamanesh N, Nelson MN, Yan G, Clarke HJ, Siddique KHM (2010) Mapping a major gene for growth habit and QTLs for ascochyta blight resistance and flowering time in a population between chickpea and Cicer reticulatum. Euphytica 173 (3):307-319.

Aukerman MJ, Sakai H (2003) Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and lts APETALA2-Like Target Genes. Plant Cell 15 (11):2730-2741.

Aung B, Gruber MY, Amyot L, Omari K, Bertrand A, Hannoufa A (2015) MicroRNA156 as a promising tool for alfalfa improvement. Plant Biotechnol J 13 (6):779-790.

Bajaj D, Upadhyaya HD, Das S, Kumar V, Gowda CL, Sharma S, Tyagi AK, Parida SK (2016) Identification of candidate genes for dissecting complex branch number trait in chickpea. Plant Sci 245:61-70.

Balasubramanian S, Sureshkumar S, Lempe J, Weigel D (2006) Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. PLoS Genet 2 (7):0980-0989.

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16 (1):37-48.

Banfield MJ, Brady RL (2000) The structure of Antirrhinum centroradialis protein (CEN) suggests a role as a kinase regulator11Edited by I. A. Wilson. Journal of Molecular Biology 297 (5):1159-1170.

Becker A, Theißen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Molecular Phylogenetics and Evolution 29 (3):464-489.

Benlloch R, Berbel A, Ali L, Gohari G, Millan T, Madueno F (2015) Genetic control of inflorescence architecture in legumes. Front Plant Sci 6:543.

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E (2015) The Arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis (New York, NY : 2000) 53 (8):474-485.

Berbel A, Ferrándiz C, Hecht V, Dalmais M, Lund OS, Sussmilch FC, Taylor SA, Bendahmane A, Ellis THN, Beltrán JP, Weller JL, Madueño F (2012) VEGETATIVE1 is essential for development of the compound inflorescence in pea. Nature Communications 3.

Berger J, Abbo S, Turner NC (2003) Ecogeography of annual wild Cicer species: The poor state of the world collection. Crop Science 43 (3):1076-1090.

Berger J, Turner NC, Buck RP Wild and cultivated Cicer species-different evolutionary paths lead to different phenological strategies that can be exploited to broaden the adaptation of chickpea (C. arietinum L.). In: *Proc. of the 4th international Crop Science Congress*, Brisbane, Australia, 26 Sept-1 Oct 2004. (www.cropscience.org.au), 26 Sept-1 Oct 2004 2004a.

Berger JD (2007) Ecogeographic and evolutionary approaches to improving adaptation of autumn-sown chickpea (Cicer arietinum L.) to terminal drought: The search for reproductive chilling tolerance. Field Crops Research 104 (1-3):112-122.

Berger JD (2014) An evolutionary perspective on the role of phenology in the specific adaptation of chickpea. Legume perspectives (3):8-11.

Berger JD, Ali M, Basu PS, Chaudhary BD, Chaturvedi SK, Deshmukh PS, Dharmaraj PS, Dwivedi SK, Gangadhar GC, Gaur PM, Kumar J, Pannu RK, Siddique KHM, Singh DN, Singh DP, Singh SJ, Turner NC, Yadava HS, Yadav SS (2006) Genotype by environment studies demonstrate the critical role of phenology in adaptation of chickpea (Cicer arietinum L.) to high and low yielding environments of India. Field Crops Research 98 (2-3):230-244.

Berger JD, Buck R, Henzell JM, Turner NC (2005) Evolution in the genus Cicer - Vernalisation response and low temperature pod set in chickpea (C. arietinum L.) and its annual wild relatives. Australian Journal of Agricultural Research 56 (11):1191-1200.

Berger JD, Kumar S, Nayyar H, Street KA, Sandhu JS, Henzell JM, Kaur J, Clarke HC (2012) Temperature-stratified screening of chickpea (Cicer arietinum L.) genetic resource collections reveals very limited reproductive chilling tolerance compared to its annual wild relatives. Field Crops Research 126:119-129.

Berger JD, Milroy SP, Turner NC, Siddique KHM, Imtiaz M, Malhotra R (2011) Chickpea evolution has selected for contrasting phenological mechanisms among different habitats. Euphytica 180 (1):1-15.

Berger JD, Turned NC, Siddique KHM, Knights EJ, Brinsmead RB, Mock I, Edmondson C, Khan TN (2004b) Genotype by environment studies across Australia reveal the importance of phenology for chickpea (Cicer arietinum L.) improvement. Australian Journal of Agricultural Research 55 (10):1071-1084.

Berger JD, Turner NC (2007) The ecology of chickpea. In: Chickpea Breeding and Management. CABI Publishing, pp 47-71

Berry G, Aitken Y (1979) Effect of photoperiod and temperature on flowering in pea (Pisum sativum L.). Funct Plant Biol 6 (6):573-587.

Berry S, Dean C (2015) Environmental perception and epigenetic memory: mechanistic insight through FLC. Plant J 83 (1):133-148.

Berry S, Hartley M, Olsson TS, Dean C, Howard M (2015) Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. eLife 4.

Beveridge CA, Kyozuka J (2010) New genes in the strigolactone-related shoot branching pathway. Curr Opin Plant Biol 13 (1):34-39.

Beveridge CA, Weller JL, Singer SR, Hofer JMI (2003) Axillary meristem development. Budding relationships between networks controlling flowering, branching, and photoperiod responsiveness. Plant Physiology 131 (3):927-934.

Blackman BK, Strasburg JL, Raduski AR, Michaels SD, Rieseberg LH (2010) The role of recently derived FT paralogs in sunflower domestication. Curr Biol 20 (7):629-635. doi: 610.1016/j.cub.2010.1001.1059. Epub 2010 Mar 1018.

Blázquez MA, Ahn JH, Weigel D (2003) A thermosensory pathway controlling flowering time in Arabidopsis thaliana. Nature Genetics 33 (2):168-171.

Blazquez MA, Green R, Nilsson O, Sussman MR, Weigel D (1998) Gibberellins promote flowering of arabidopsis by activating the LEAFY promoter. Plant Cell 10 (5):791-800.

Bohlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. Science 312 (5776):1040-1043.

Bonfil D, Lichtenzveig J, Shai I, Lerner A, Tam S, Abbo S (2006a) Associations between earliness, Ascochyta response, and grain yield in chickpea. Crop Pasture Sci 57 (4):465-470.

Bonfil DJ, Lichtenzveig J, Shai I, Lerner A, Tam S, Abbo S (2006b) Associations between earliness, Ascochyta response, and grain yield in chickpea. Australian Journal of Agricultural Research 57 (4):465-470.

Bradley D, Ratcliffe O, Vincent C, Carpenter R, Coen E (1997) Inflorescence Commitment and Architecture in Arabidopsis. Science 275 (5296):80-83.

Bratzel F, Turck F (2015) Molecular memories in the regulation of seasonal flowering: From competence to cessation. Genome Biol 16 (1).

Braun N, de Saint Germain A, Pillot JP, Boutet-Mercey S, Dalmais M, Antoniadi I, Li X, Maia-Grondard A, Le Signor C, Bouteiller N, Luo D, Bendahmane A, Turnbull C, Rameau C (2012) The pea TCP transcription factor PsBRC1 acts downstream of Strigolactones to control shoot branching. Plant Physiol 158 (1):225-238.

Bull SE, Alder A, Barsan C, Kohler M, Hennig L, Gruissem W, Vanderschuren H (2017) FLOWERING LOCUS T Triggers Early and Fertile Flowering in Glasshouse Cassava (Manihot esculenta Crantz). Plants (Basel, Switzerland) 6 (2).

Cai Y, Chen X, Xie K, Xing Q, Wu Y, Li J, Du C, Sun Z, Guo Z (2014) Dlf1, a WRKY Transcription Factor, Is Involved in the Control of Flowering Time and Plant Height in Rice. Plos One 9 (7):e102529.

Campoli C, Pankin A, Drosse B, Casao CM, Davis SJ, von Korff M (2013) HvLUX1 is a candidate gene underlying the early maturity 10 locus in barley: phylogeny, diversity, and interactions with the circadian clock and photoperiodic pathways. New Phytol 199 (4):1045-1059.

Campos-Rivero G, Osorio-Montalvo P, Sanchez-Borges R, Us-Camas R, Duarte-Ake F, De-la-Pena C (2017) Plant hormone signaling in flowering: An epigenetic point of view. J Plant Physiol 214:16-27.

Cao D, Li Y, Lu S, Wang J, Nan H, Li X, Shi D, Fang C, Zhai H, Yuan X, Anai T, Xia Z, Liu B, Kong F (2015a) GmCOL1a and GmCOL1b Function as Flowering Repressors in Soybean Under Long-Day Conditions. Plant Cell Physiol 56 (12):2409-2422.

Cao D, Li Y, Wang J, Nan H, Wang Y, Lu S, Jiang Q, Li X, Shi D, Fang C, Yuan X, Zhao X, Li X, Liu B, Kong F (2015b) GmmiR156b overexpression delays flowering time in soybean. Plant Mol Biol 89 (4-5):353-363.

Cao S, Kumimoto RW, Gnesutta N, Calogero AM, Mantovani R, Holt BF (2014) A Distal CCAAT/NUCLEAR FACTOR Y Complex Promotes Chromatin Looping at the

FLOWERING LOCUS T Promoter and Regulates the Timing of Flowering in Arabidopsis. The Plant Cell Online.

Carmona MJ, Calonje M, Martínez-Zapater JM (2007) The FT/TFL1 gene family in grapevine. Plant Mol Biol 63 (5):637-650.

Chae HS, Faure F, Kieber JJ (2003) The eto1, eto2, and eto3 Mutations and Cytokinin Treatment Increase Ethylene Biosynthesis in Arabidopsis by Increasing the Stability of ACS Protein. The Plant Cell 15 (2):545-559.

Chaurasia AK, Patil HB, Krishna B, Subramaniam VR, Sane PV, Sane AP (2017) Flowering time in banana (Musa spp.), a day neutral plant, is controlled by at least three FLOWERING LOCUS T homologues. Sci Rep 7 (1):5935.

Cheng JZ, Zhou YP, Lv TX, Xie CP, Tian CE (2017) Research progress on the autonomous flowering time pathway in Arabidopsis. Physiology and molecular biology of plants : an international journal of functional plant biology 23 (3):477-485.

Cheng XF, Wang ZY (2005) Overexpression of COL9, a CONSTANS-LIKE gene, delays flowering by reducing expression of CO and FT in Arabidopsis thaliana. Plant J 43 (5):758-768.

Cho S, Kumar J, Shultz JL, Anupama K, Tefera F, Muehlbauer FJ (2002) Mapping genes for double podding and other morphological traits in chickpea. Euphytica 128 (2):285-292.

Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND, Cook DR (2004) Estimating genome conservation between crop and model legume species. Proceedings of the National Academy of Sciences of the United States of America 101 (43):15289-15294.

Clancy S (2008) RNA Functions. Nature Education 1(1):102.

Clarke HJ, Siddique KHM (2004) Response of chickpea genotypes to low temperature stress during reproductive development. Field Crops Research 90 (2-3):323-334.

Cobos MJ, Fernández MJ, Rubio J, Kharrat M, Moreno MT, Gil J, Millán T (2005) A linkage map of chickpea (Cicer arietinum L.) based on populations from Kabuli × Desi crosses: Location of genes for resistance to fusarium wilt race 0. Theoretical and Applied Genetics 110 (7):1347-1353.

Cobos MJ, Rubio J, Fernández-Romero MD, Garza R, Moreno MT, Millán T, Gil J (2007) Genetic analysis of seed size, yield and days to flowering in a chickpea recombinant inbred line population derived from a Kabuli × Desi cross. Annals of Applied Biology 151 (1):33-42.

Cobos MJ, Winter P, Kharrat M, Cubero JI, Gil J, Millan T, Rubio J (2009) Genetic analysis of agronomic traits in a wide cross of chickpea. Field Crops Research 111 (1-2):130-136.

Conti L (2017) Hormonal control of the floral transition: Can one catch them all? Dev Biol 430 (2):288-301.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science 316 (5827):1030-1033.

Coustham V, Li P, Strange A, Lister C, Song J, Dean C (2012) Quantitative modulation of polycomb silencing underlies natural variation in vernalization. Science 337 (6094):584-587.

Croser JS, Ahmad F, Clarke HJ, Siddique KHM (2003a) Utilisation of wild Cicer in chickpea improvement - Progress, constraints, and prospects. Australian Journal of Agricultural Research 54 (5):429-444.

Croser JS, Clarke HJ, Siddique KHM, Khan TN (2003b) Low-temperature stress: Implications for chickpea (Cicer arietinum L.) improvement. Crit Rev Plant Sci 22 (2):185-219.

Cruz-Izquierdo S, Avila CM, Satovic Z, Palomino C, Gutierrez N, Ellwood SR, Phan HT, Cubero JI, Torres AM (2012) Comparative genomics to bridge Vicia faba with model and closely-related legume species: stability of QTLs for flowering and yield-related traits. Theor Appl Genet 125 (8):1767-1782.

Cui X, Cao X (2014) Epigenetic regulation and functional exaptation of transposable elements in higher plants. Current Opinion in Plant Biology 21 (Supplement C):83-88.

Cui XK, Cao XF (2015) Overview of the function of transposable elements in higher plants. Progress in Biochemistry and Biophysics 42 (11):1033-1046.

D'Aloia M, Bonhomme D, Bouche F, Tamseddak K, Ormenese S, Torti S, Coupland G, Perilleux C (2011) Cytokinin promotes flowering of Arabidopsis via transcriptional activation of the FT paralogue TSF. Plant J 65 (6):972-979.

Daba K, Deokar A, Banniza S, Warkentin TD, Tar'an B (2016a) QTL mapping of early flowering and resistance to ascochyta blight in chickpea. Genome 59 (6):413-425.

Daba K, Tar'an B, Warkentin TD (2015) Flowering response of diverse chickpea (Cicer arietinum L.) accessions to photoperiod. Genet Resour Crop Evol:1-12.

Daba K, Taran B, Bueckert R, Warkentin TD (2016b) Effect of temperature and photoperiod on time to flowering in chickpea. Crop Science 56 (1):200-208.

Daba K, Warkentin TD, Bueckert R, Todd CD, Tar'an B (2016c) Determination of Photoperiod-Sensitive Phase in Chickpea (Cicer arietinum L.). Front Plant Sci 7:478.

Danilevskaya ON, Meng X, Hou Z, Ananiev EV, Simmons CR (2008) A genomic and expression compendium of the expanded PEBP gene family from maize. Plant Physiol 146 (1):250-264. Epub 2007 Nov 2009.

Das S, Upadhyaya HD, Bajaj D, Kujur A, Badoni S, Laxmi, Kumar V, Tripathi S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015a) Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes 22 (3):193-203.

Das S, Upadhyaya HD, Srivastava R, Bajaj D, Gowda CL, Sharma S, Singh S, Tyagi AK, Parida SK (2015b) Genome-wide insertion-deletion (InDel) marker discovery and genotyping for genomics-assisted breeding applications in chickpea. DNA research : an international journal for rapid publication of reports on genes and genomes 22 (5):377-386.

Datta S, Hettiarachchi GH, Deng XW, Holm M (2006) Arabidopsis CONSTANS-LIKE3 is a positive regulator of red light signaling and root growth. Plant Cell 18 (1):70-84.

Davies PJ (1995) Plant Hormones-Physiology, Biochemistry and Molecular Biology. Springer, Dordrecht,

Deng XW, Xu D, Zhu D (2016) The role of COP1 in repression of photoperiodic flowering. F1000 Res 5.

Dennis ES, Peacock WJ (2009) Vernalization in cereals. Journal of biology 8 (6):57.

Deokar AA, Ramsay L, Sharpe AG, Diapari M, Sindhu A, Bett K, Warkentin TD, Tar'an B (2014) Genome wide SNP identification in chickpea for use in development of a high density genetic map and improvement of chickpea reference genome assembly. BMC Genomics 15 (1):1471-2164.

Devasirvatham V, Gaur PM, Raju TN, Trethowan RM, Tan DKY (2015) Field response of chickpea (Cicer arietinum L.) to high temperature. Field Crops Research 172:59-71.

Devasirvatham V, Tan DKY, Gaur PM, Raju TN, Trethowan RM (2012) High temperature tolerance in chickpea and its implications for plant improvement. Crop Pasture Sci 63 (5):419-422.

Distelfeld A, Li C, Dubcovsky J (2009) Regulation of flowering in temperate cereals. Curr Opin Plant Biol 12 (2):178-184.

Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF (2004) The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. Curr Biol 14 (21):1935-1940.

Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature 386 (6624):485-488.

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127 (7):1309-1321.

Domagalska MA, Leyser O (2011) Signal integration in the control of shoot branching. Nat Rev Mol Cell Biol 12 (4):211-221.

Doyle MR, Davis SJ, Bastow RM, McWatters HG, Kozma-Bognar L, Nagy F, Millar AJ, Amasino RM (2002) The ELF4 gene controls circadian rhythms and flowering time in Arabidopsis thaliana. Nature 419 (6902):74-77.

Dreni L, Zhang D (2016) Flower development: the evolutionary history and functions of the AGL6 subfamily MADS-box genes. J Exp Bot 67 (6):1625-1638.

Durfee T, Roe JL, Sessions RA, Inouye C, Serikawa K, Feldmann KA, Weigel D, Zambryski PC (2003) The F-box-containing protein UFO and AGAMOUS participate in antagonistic

pathways governing early petal development in Arabidopsis. Proc Natl Acad Sci U S A 100 (14):8571-8576.

Dusunceli F, Wood JA, Gupta A, Yadav M, Yadav SS (2007) International trade. In: Chickpea Breeding and Management. CABI Publishing, pp 555-575

Eckardt NA (2007) Two Tales of Chromatin Remodeling Converge on HUB1. The Plant Cell 19 (2):391-393.

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5 (1):113.

Ellis RH, Lawn RJ, Summerfield RJ, Qi A, Roberts EH, Chay PM, Brouwer JB, Rose JL, Yeates SJ, Sandover S (1994) Towards a reliable prediction of time to flowering in six annual crops. V. Chickpea (*Cicer arietinum*). Experimental Agriculture 30:271-282.

Erskine W, Ellis RH, Summerfield RJ, Roberts EH, Hussain A (1990) Characterization of responses to temperature and photoperiod for time to flowering in a world lentil collection. Theoretical and Applied Genetics 80 (2):193-199.

Ezer D, Jung JH, Lan H, Biswas S, Gregoire L, Box MS, Charoensawan V, Cortijo S, Lai X, Stockle D, Zubieta C, Jaeger KE, Wigge PA (2017) The evening complex coordinates environmental and endogenous signals in Arabidopsis. Nature plants 3:17087.

Ezzat K, Reza T, Mehrab K, Abbas S (2015) A linkage map of chickpea (Cicer arietinum L.) based on population from ILC3279×ILC588 crosses: Location of genes for time to flowering, seed size and plant height. Genetika 47 (1):253-263.

Fang X, Turner NC, Yan G, Li F, Siddique KH (2010) Flower numbers, pod production, pollen viability, and pistil function are reduced and flower and pod abortion increased in chickpea (Cicer arietinum L.) under terminal drought. J Exp Bot 61 (2):335-345.

Feng W, Michaels SD (2011) Dual roles for FY in the regulation of FLC. Plant Signaling & Behavior 6 (5):703-705.

Fernandez V, Takahashi Y, Le Gourrierec J, Coupland G (2016) Photoperiodic and thermosensory pathways interact through CONSTANS to promote flowering at high temperature under short days. Plant J 86 (5):426-440.

Ferrandiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development 127 (4):725-734.

Finlayson SA (2007) Arabidopsis Teosinte Branched1-like 1 regulates axillary bud outgrowth and is homologous to monocot Teosinte Branched1. Plant Cell Physiol 48 (5):667-677.

Finnegan DJ (2012) Retrotransposons. Curr Biol 22 (11):R432-R437.

Fornara F, Panigrahi KCS, Gissot L, Sauerbrunn N, Rühl M, Jarillo JA, Coupland G (2009) Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. Dev Cell 17 (1):75-86.

Foucher F, Morin J, Courtiade J, Cadioux S, Ellis N, Banfield MJ, Rameau C (2003) DETERMINATE and LATE FLOWERING are two TERMINAL FLOWER1/CENTRORADIALIS homologs that control two distinct phases of flowering initiation and development in pea. Plant Cell 15 (11):2742-2754.

Galindo-González L, Mhiri C, Deyholos MK, Grandbastien MA (2017) LTR-retrotransposons in plants: Engines of evolution. Gene 626:14-25.

Galvão VC, Schmid M (2014) Regulation of flowering by endogenous signals. The Molecular Genetics of Floral Transition and Flower Development 72:63-102.

Garg R, Patel RK, Tyagi AK, Jain M (2011) De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification. DNA research 18 (1):53-63.

Gaur PM, Jukanti AK, Varshney RK (2012) Impact of Genomic Technologies on Chickpea Breeding Strategies. Agronomy 2 (3):199-221.

Gaur PM, Kumar J, Gowda CLL, Pande S, Siddique KHM, Khan TN, Warkentin TD, Chaturvedi SK, Than AM, Ketema D (2008) Breeding chickpea for early phenology: perspectives, progress and prospects.

Gaur PM, Samineni S, Tripathi S, Varshney RK, Gowda CLL (2014a) Allelic relationships of flowering time genes in chickpea. Euphytica.

Gaur PM, Thudi M, Samineni S, Varshney RK (2014b) Advances in Chickpea Genomics. In: Legumes in the Omic Era. Springer, pp 73-94

Gil J, Cubero J (1993) Inheritance of seed coat thickness in chickpea (Cicer arietinum L.) and its evolutionary implications. Plant Breeding 111 (3):257-260.

Gocal GFW, Sheldon CC, Gubler F, Moritz T, Bagnall DJ, MacMillan CP, Li SF, Parish RW, Dennis ES, Weigel D, King RW (2001) GAMYB-like genes, flowering, and gibberellin signaling in Arabidopsis. Plant Physiology 127 (4):1682-1693.

Gondo T, Sato S, Okumura K, Tabata S, Akashi R, Isobe S (2007) Quantitative trait locus analysis of multiple agronomic traits in the model legume Lotus japonicus. Genome 50 (7):627-637.

González A, Martín I, Ayerbe L (1999) Barley yield in water-stress conditions. The influence of precocity, osmotic adjustment and stomatal conductance. Field Crops Research 62 (1):23-34.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40 (Database issue):D1178-D1186.

Gowda SJ, Radhika P, Mhase LB, Jamadagni BM, Gupta VS, Kadoo NY (2011) Mapping of QTLs governing agronomic and yield traits in chickpea. Journal of applied genetics 52 (1):9-21.

Gramzow L, Theißen G (2013) Phylogenomics of MADS-box genes in plants - Two opposing life styles in one gene family. Biology 2 (3):1150-1164.

Greenham K, McClung CR (2015) Integrating circadian dynamics with physiological processes in plants. Nat Rev Genet 16 (10):598-610.

Greenup A, Peacock WJ, Dennis ES, Trevaskis B (2009) The molecular biology of seasonal flowering-responses in Arabidopsis and the cereals. Ann Bot 103 (8):1165-1172.

Griffiths S, Dunford RP, Coupland G, Laurie DA (2003) The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis. Plant Physiol 131 (4):1855-1867.

Gu X, Jiang D, Wang Y, Bachmair A, He Y (2009) Repression of the floral transition via histone H2B monoubiquitination. The Plant Journal 57 (3):522-533.

Gujaria-Verma N, Vail SL, Carrasquilla-Garcia N, Penmetsa RV, Cook DR, Farmer AD, Vandenberg A, Bett KE (2014) Genetic mapping of legume orthologs reveals high conservation of synteny between lentil species and the sequenced genomes of Medicago and chickpea. Front Plant Sci 5:676.

Gujaria N, Kumar A, Dauthal P, Dubey A, Hiremath P, Bhanu Prakash A, Farmer A, Bhide M, Shah T, Gaur PM, Upadhyaya HD, Bhatia S, Cook DR, May GD, Varshney RK (2011) Development and use of genic molecular markers (GMMs) for construction of a transcript map of chickpea (Cicer arietinum L.). TAG Theoretical and Applied Genetics Theoretische Und Angewandte Genetik 122 (8):1577-1589.

Gumber RK, Sarvjeet S (1996) Genetics of flowering time in chickpea: a preliminary report. Crop Improvement 23:295-296.

Hamwieh A, Imtiaz M, Malhotra R (2013a) Multi-environment QTL analyses for drought-related traits in a recombinant inbred population of chickpea (Cicer arientinum L.). Theoretical and applied genetics 126 (4):1025-1038.

Hamwieh A, Imtiaz M, Malhotra RS (2013b) Multi-environment QTL analyses for drought-related traits in a recombinant inbred population of chickpea (Cicer arientinum L.). Theoretical and Applied Genetics 126 (4):1025-1038.

Hanzawa Y, Money T, Bradley D (2005) A single amino acid converts a repressor to an activator of flowering. Proceedings of the National Academy of Sciences of the United States of America 102 (21):7748-7753.

Hassidim M, Harir Y, Yakir E, Kron I, Green RM (2009) Over-expression of CONSTANS-LIKE 5 can induce flowering in short-day grown Arabidopsis. Planta 230 (3):481-491.

Hayama R, Sarid-Krebs L, Richter R, Fernández V, Jang S, Coupland G (2017) PSEUDO RESPONSE REGULATORs stabilize CONSTANS protein to promote flowering in response to day length. The EMBO Journal.

Hazen SP, Schultz TF, Pruneda-Paz JL, Borevitz JO, Ecker JR, Kay SA (2005) LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. Proc Natl Acad Sci U S A 102 (29):10387-10392.

He Y (2012) Chromatin regulation of flowering. Trends Plant Sci 17 (9):556-562.

Hecht V, Foucher F, Ferrándiz C, Macknight R, Navarro C, Morin J, Vardy ME, Ellis N, Beltrán JP, Rameau C, Weller JL (2005) Conservation of Arabidopsis flowering genes in model legumes. Plant Physiology 137 (4):1420-1434.

Hecht V, Knowles CL, Vander Schoor JK, Liew LC, Jones SE, Lambert MJM, Weller JL (2007a) Pea *LATE BLOOMER1* is a *GIGANTEA* ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. Plant Physiol 144 (2):648-661.

Hecht V, Knowles CL, Vander Schoor JK, Lim CL, Jones SE, Lambert MJM, Weller JL (2007b) Pea Late Bloomer1 is a Gigantea ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. Plant Physiology 144 (2):648-661.

Hecht V, Laurie RE, Vander Schoor JK, Ridge S, Knowles CL, Liew LC, Sussmilch FC, Murfet IC, Macknight RC, Weller JL (2011) The pea GIGAS gene is a FLOWERING LOCUS T homolog necessary for graft-transmissible specification of flowering but not for responsiveness to photoperiod. Plant Cell 23 (1):147-161.

Hedden P, Phillips AL (2000) Gibberellin metabolism: new insights revealed by the genes. Trends in Plant Science 5 (12):523-530.

Hegde VS (2010) Genetics of flowering time in chickpea in a semi-arid environment. Plant Breeding 129 (6):683-687.

Helfer A, Nusinow DA, Chow BY, Gehrke AR, Bulyk ML, Kay SA (2011) LUX ARRHYTHMO Encodes a Night Time Repressor of Circadian Gene Expression in the Arabidopsis Core Clock. Current biology : CB 21 (2):126-133.

Helliwell CA, Wood CC, Robertson M, James Peacock W, Dennis ES (2006) The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. Plant J 46 (2):183-192.

Hemming MN, Peacock WJ, Dennis ES, Trevaskis B (2008) Low-temperature and daylength cues are integrated to regulate FLOWERING LOCUS T in barley. Plant Physiol 147 (1):355-366.

Hiraoka K, Yamaguchi A, Abe M, Araki T (2013) The Florigen Genes FT and TSF Modulate Lateral Shoot Outgrowth in Arabidopsis thaliana. Plant Cell Physiol 54 (3):352-368.

Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H, Tuteja R, Kumar A, BhanuPrakash A, Mulaosmanovic B, Gujaria N, Krishnamurthy L, Gaur PM, KaviKishor PB, Shah T, Srinivasan R, Lohse M, Xiao Y, Town CD, Cook DR, May GD, Varshney RK (2011) Large-scale transcriptome analysis in chickpea (Cicer arietinum L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. Plant Biotechnol J 9 (8):922-931.

Hiremath PJ, Kumar A, Penmetsa RV, Farmer A, Schlueter JA, Chamarthi SK, Whaley AM, Carrasquilla-Garcia N, Gaur PM, Upadhyaya HD, Kavi Kishor PB, Shah TM, Cook DR, Varshney RK (2012) Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. Plant Biotechnol J 10 (6):716-732.

Hirsch CD, Springer NM (2017) Transposable element influences on gene expression in plants. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1860 (1):157-165.

Hisamatsu T, King RW (2008) The nature of floral signals in Arabidopsis. II. Roles for FLOWERING LOCUS T (FT) and gibberellin. Journal of Experimental Botany 59 (14):3821-3829.

Ho WWH, Weigel D (2014) Structural features determining flower-promoting activity of Arabidopsis FLOWERING LOCUS T. Plant Cell 26 (2):552-564.

Hori K, Matsubara K, Uga Y, Yano M (2016) A novel Tos17 insertion upstream of Hd1 alters flowering time in rice. Plant Breeding 135 (5):588-592.

Hossain S, Ford R, McNeil D, Pittock C, Panozzo JF (2010) Development of a selection tool for seed shape and QTL analysis of seed shape with other morphological traits for selective breeding in chickpea (Cicer arietinum L.). Aust J Crop Sci 4 (4):278-288.

Hovav R, Upadhyaya KC, Beharav A, Abbo S (2003) Major flowering time gene and polygene effects on chickpea seed weight. Plant Breeding 122 (6):539-541.

Hsu C-Y, Liu Y, Luthe DS, Yuceer C (2006) Poplar FT2 Shortens the Juvenile Phase and Promotes Seasonal Flowering. The Plant Cell 18 (8):1846-1861.

Hsu HF, Huang CH, Chou LT, Yang CH (2003) Ectopic expression of an orchid (Oncidium Gower Ramsey) AGL6-like gene promotes flowering by activating flowering time genes in Arabidopsis thaliana. Plant Cell Physiol 44 (8):783-794.

Huang H, Nusinow DA (2016) Into the Evening: Complex Interactions in the Arabidopsis Circadian Clock. Trends Genet 32 (10):674-686.

Huang S, Raman AS, Ream JE, Fujiwara H, Cerny RE, Brown SM (1998) Overexpression of 20-Oxidase Confers a Gibberellin-Overproduction Phenotype in Arabidopsis. Plant Physiology 118 (3):773-781.

Huang XQ, Ding J, Effgen S, Turck F, Koornneef M (2013) Multiple loci and genetic interactions involving flowering time genes regulate stem branching among natural variants of Arabidopsis. New Phytol 199 (3):843-857.

Hubbard L, McSteen P, Doebley J, Hake S (2002) Expression Patterns and Mutant Phenotype of *teosinte branched1* Correlate With Growth Suppression in Maize and Teosinte. Genetics 162 (4):1927-1935.

Huijser P, Schmid M (2011) The control of developmental phase transitions in plants. Development 138 (19):4117-4129.

Hussien A, Tavakol E, Horner DS, Muñoz-Amatriaín M, Muehlbauer GJ, Rossini L (2014) Genetics of Tillering in Rice and Barley. The Plant Genome 7 (1).

Huyghe C (1998) Genetics and genetic modifications of plant architecture in grain legumes: a review. Agronomie 18 (5-6):383-411.

Iliadis C (2001) Evaluation of six chickpea varieties for seed yield under autumn and spring sowing. Journal of Agricultural Science 137 (4):439-444.

Imaizumi T, Tran HG, Swartz TE, Briggs WR, Kay SA (2003) FKF1 is essential for photoperiodic-specific light signalling in Arabidopsis. Nature 426 (6964):302-306.

Inouye M (1988) Antisense RNA: its functions and applications in gene regulation — a review. Gene 72 (1):25-34.

Iruela M, Castro P, Rubio J, Cubero JI, Jacinto C, Millán T, Gil J (2007) Validation of a QTL for resistance to ascochyta blight linked to resistance to fusarium wilt race 5 in chickpea (Cicer arietinum L.). In: Tivoli B, Baranger A, Muehlbauer FJ, Cooke BM (eds) Ascochyta blights of grain legumes. Springer Netherlands, Dordrecht, pp 29-37. doi:10.1007/978-1-4020-6065-6_4

Iruela M, Rubio J, Cubero JI, Gil J, Millán T (2002) Phylogenetic analysis in the genus Cicer and cultivated chickpea using RAPD and ISSR markers. Theoretical and Applied Genetics 104 (4):643-651.

Irwin JA, Soumpourou E, Lister C, Ligthart JD, Kennedy S, Dean C (2016) Nucleotide polymorphism affecting FLC expression underpins heading date variation in horticultural brassicas. Plant J 87 (6):597-605.

Jain M, Misra G, Patel RK, Priya P, Jhanwar S, Khan AW, Shah N, Singh VK, Garg R, Jeena G, Yadav M, Kant C, Sharma P, Yadav G, Bhatia S, Tyagi AK, Chattopadhyay D (2013) A draft genome sequence of the pulse crop chickpea (Cicer arietinum L.). Plant J 74 (5):715-729.

Jamalabadi JG, Saidi A, Karami E, Kharkesh M, Talebi R (2013) Molecular mapping and characterization of genes governing time to flowering, seed weight, and plant height in an intraspecific genetic linkage map of chickpea (Cicer arietinum). Biochemical Genetics 51 (5-6):387-397.

Janssen BJ, Drummond RS, Snowden KC (2014) Regulation of axillary shoot development. Curr Opin Plant Biol 17:28-35.

Jaudal M, Yeoh CC, Zhang LL, Stockum C, Mysore KS, Ratet P, Putterill J (2013) Retroelement insertions at the Medicago FTa1 locus in spring mutants eliminate vernalisation but not long-day requirements for early flowering. Plant J 76 (4):580-591.

Jaudal M, Zhang L, Che C, Putterill J (2015) Three Medicago MtFUL genes have distinct and overlapping expression patterns during vegetative and reproductive development and 35S:MtFULb accelerates flowering and causes a terminal flower phenotype in Arabidopsis. Frontiers in genetics 6:50.

Javadi F, Wojciechowski MF, Yamaguchi H (2007) Geographical diversification of the genus Cicer (Leguminosae: Papilionoideae) inferred from molecular phylogenetic analyses of chloroplast and nuclear DNA sequences. Botanical Journal of the Linnean Society 154 (2):175-186.

Jayakumar P, Gossen BD, Gan YT, Warkentin TD, Banniza S (2005) Ascochyta blight of chickpea: Infection and host resistance mechanisms. Canadian Journal of Plant Pathology 27 (4):499-509.

Jeong HL, Seong JY, Soo HP, Hwang I, Jong SL, Ji HA (2007) Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. Genes and Development 21 (4):397-402.

Jern P, Coffin JM (2008) Effects of retroviruses on host genome function. Annual review of genetics 42:709-732.

Jhanwar S, Priya P, Garg R, Parida SK, Tyagi AK, Jain M (2012) Transcriptome sequencing of wild chickpea as a rich resource for marker development. Plant Biotechnol J 10 (6):690-702.

Johansen C, Singh DN, Krishnamurthy L, Saxena NP, Chauhan YS, Kumar Rao JVDK (1997) Options for alleviating moisture stress in pulse crops. In: Asthana AN, Ali M (eds) Recent Advances in Pulses Research. Indian Institute of Pulses Research, Indian Society of Pulses Research and Development, Kanpur, Uttar Pradesh, India, pp 425-442

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C (2000) Molecular Analysis of *FRIGIDA*, a Major Determinant of Natural Variation in *Arabidopsis* Flowering Time. Science 290 (5490):344-347.

Jukanti AK, Gaur PM, Gowda CLL, Chibbar RN (2012) Nutritional quality and health benefits of chickpea (Cicer arietinum L.): A review. Br J Nutr 108 (SUPPL. 1).

Julier B, Huguet T, Chardon F, Ayadi R, Pierre JB, Prosperi JM, Barre P, Huyghe C (2007) Identification of quantitative trait loci influencing aerial morphogenesis in the model legume Medicago truncatula. Theor Appl Genet 114 (8):1391-1406.

Jung C-H, Wong CE, Singh MB, Bhalla PL (2012) Comparative Genomic Analysis of Soybean Flowering Genes. Plos One 7 (6):e38250.

Jung JH, Seo YH, Pil JS, Reyes JL, Yun J, Chua NH, Park CM (2007) The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in Arabidopsis. Plant Cell 19 (9):2736-2748.

Kaikkonen MU, Lam MTY, Glass CK (2011) Non-coding RNAs as regulators of gene expression and epigenetics. Cardiovascular Research 90 (3):430-440.

Kanouni H, Taleei A, Okhovat M (2011) Ascochyta blight (Ascochyta rabiei (Pass.) Lab.) of Chickpea (Cicer arietinum L.): Breeding Strategies for Resistance. International Journal of Plant Breeding and Genetics 5:1-22.

Karlgren A, Gyllenstrand N, Kallman T, Sundstrom JF, Moore D, Lascoux M, Lagercrantz U (2011) Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. Plant Physiol 156 (4):1967-1977.

Kassie M, Shiferaw B, Asfaw S, Abate T, Muricho G, Ferede S, Eshete M, Assefa K (2009) Current situation and future outlooks of the chickpea sub-sector in Ethiopia. ICRISAT and EIAR (http://www icrisat org/tropicallegumesII/pdfs/Current_Situation pdf).

Katoh K, Misawa K, Kuma K-i, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30 (14):3059-3066.

Kazan K, Muehlbauer FJ, Weeden NE, Ladizinsky G (1993) Inheritance and linkage relationships of morphological and isozyme loci in chickpea (Cicer arietinum L.). Theoretical and Applied Genetics 86 (4):417-426.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012a) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12):1647-1649.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A (2012b) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12):1647-1649.

Kebrom TH, Spielmeyer W, Finnegan EJ (2013) Grasses provide new insights into regulation of shoot branching. Trends in Plant Science 18 (1):41-48.

Kempin SA, Savidge B, Yanofsky MF (1995) Molecular basis of the cauliflower phenotype in Arabidopsis. Science 267 (5197):522-525.

Kikuchi R, Kawahigashi H, Ando T, Tonooka T, Handa H (2009) Molecular and Functional Characterization of PEBP Genes in Barley Reveal the Diversification of Their Roles in Flowering. Plant Physiology 149 (3):1341-1353.

Kim DH, Doyle MR, Sung S, Amasino RM (2009) Vernalization: Winter and the timing of flowering in plants. Annual Review of Cell and Developmental Biology, vol 25. doi:10.1146/annurev.cellbio.042308.113411

Kim J, Geng R, Gallenstein RA, Somers DE (2013a) The F-box protein ZEITLUPE controls stability and nucleocytoplasmic partitioning of GIGANTEA. Development 140 (19):4060-4069.

Kim MY, Kang YJ, Lee T, Lee SH (2013b) Divergence of flowering-related genes in three legume species. Plant Genome 6 (3):4.

Kim MY, Shin JH, Kang YJ, Shim SR, Lee SH (2012) Divergence of flowering genes in soybean. Journal of Biosciences 37 (5):857-870.

Kim SY, Yu X, Michaels SD (2008) Regulation of CONSTANS and FLOWERING LOCUS T expression in response to changing light quality. Plant Physiology 148 (1):269-279.

Kim WY, Fujiwara S, Suh SS, Kim J, Kim Y, Han L, David K, Putterill J, Nam HG, Somers DE (2007) ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature 449 (7160):356-360.

Kipreos ET, Pagano M (2000) The F-box protein family. Genome Biol 1 (5):reviews3002.3001-reviews3002.3007.

Kloosterman B, Abelenda JA, Gomez MdMC, Oortwijn M, de Boer JM, Kowitwanich K, Horvath BM, van Eck HJ, Smaczniak C, Prat S, Visser RGF, Bachem CWB (2013) Naturally occurring allele diversity allows potato cultivation in northern latitudes. Nature 495 (7440):246-250.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T (1999) A pair of related genes with antagonistic roles in mediating flowering signals. Science 286 (5446):1960-1962.

Komiya R, Ikegami A, Tamaki S, Yokoi S, Shimamoto K (2008) *Hd3a* and *RFT1* are essential for flowering in rice. Development 135 (4):767-774.

Kong F, Liu B, Xia Z, Sato S, Kim BM, Watanabe S, Yamada T, Tabata S, Kanazawa A, Harada K, Abe J (2010) Two Coordinately Regulated Homologs of FLOWERING LOCUS T Are Involved in the Control of Photoperiodic Flowering in Soybean. Plant Physiology 154 (3):1220-1231.

Kosambi DD (1943) The estimation of map distances from recombination values. Annals of Eugenics 12 (1):172-175.

Kotoda N, Hayashi H, Suzuki M, Igarashi M, Hatsuyama Y, Kidou S-i, Igasaki T, Nishiguchi M, Yano K, Shimizu T, Takahashi S, Iwanami H, Moriya S, Abe K (2010) Molecular Characterization of FLOWERING LOCUS T-Like Genes of Apple (Malus × domestica Borkh.). Plant Cell Physiol 51 (4):561-575.

Krishnakumar V, Kim M, Rosen BD, Karamycheva S, Bidwell SL, Tang H, Town CD (2015) MTGD: The Medicago truncatula genome database. Plant Cell Physiol 56 (1):e1.

Książkiewicz M, Rychel S, Nelson MN, Wyrwa K, Naganowska B, Wolko B (2016) Expansion of the phosphatidylethanolamine binding protein family in legumes: a case study of Lupinus angustifolius L. FLOWERING LOCUS T homologs, LanFTc1 and LanFTc2. BMC Genomics 17 (1):820.

Kumar J, Abbo S (2001) Genetics of flowering time in chickpea and its bearing on productivity in semiarid environments. vol 72.

Kumar J, Dusunceli F, Knights EJ, Materne M, Warkentin T, Chen W, Gaur PM, Bejiga G, Yadav SS, Satyanarayana A, Rahman MM, Yadav M (2007) Chickpea farmers. In: Chickpea Breeding and Management. CABI Publishing, pp 602-616

Kumar J, Van Rheenen HA (2000) A major gene for time of flowering in chickpea. Journal of Heredity 91 (1):67-68.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33 (7):1870-1874.

Kumar SV, Lucyshyn D, Jaeger KE, Alós E, Alvey E, Harberd NP, Wigge PA (2012) Transcription factor PIF4 controls the thermosensory activation of flowering. Nature 484 (7393):242-245.

Kumar SV, Wigge PA (2010) H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in Arabidopsis. Cell 140 (1):136-147.

Kwak M, Velasco D, Gepts P (2008) Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris). Journal of Heredity 99 (3):283-291.

Labdi M, Robertson LD, Singh KB, Charrier A (1996) Genetic diversity and phylogenetic relationships among the annual Cicer species as revealed by isozyme polymorphism. Euphytica 88 (3):181-188.

Ladizinsky G, Adler A (1976) Genetic relationships among the annual species of Cicer L. Theoretical and Applied Genetics 48 (4):197-203.

Lagercrantz U, Axelsson T (2000) Rapid evolution of the family of CONSTANS LIKE genes in plants. Mol Biol Evol 17 (10):1499-1507.

Lagunes Espinoza LC, Huguet T, Julier B (2012) Multi-population QTL detection for aerial morphogenetic traits in the model legume Medicago truncatula. Theoretical and Applied Genetics 124 (4):739-754.

Lau NC, Lai EC (2005) Diverse roles for RNA in gene regulation. Genome Biol 6 (4):315-315.

Laurie RE, Diwadkar P, Jaudal M, Zhang LL, Hecht V, Wen JQ, Tadege M, Mysore KS, Putterill J, Weller JL, Macknight RC (2011) The Medicago FLOWERING LOCUS T Homolog, MtFTa1, Is a Key Regulator of Flowering Time. Plant Physiology 156 (4):2207-2224.

Ledger S, Strayer C, Ashton F, Kay SA, Putterill J (2001) Analysis of the function of two circadian-regulated CONSTANS-LIKE genes. Plant J 26 (1):14-22.

Leduc N, Roman H, Barbier F, Péron T, Huché-Thélier L, Lothier J, Demotes-Mainard S, Sakr S (2014) Light Signaling in Bud Outgrowth and Branching in Plants. Plants 3 (2):223.

Lee C, Yu D, Choi H-K, Kim RW (2017) Reconstruction of a composite comparative map composed of ten legume genomes. Genes & Genomics 39 (1):111-119.

Lee H, Suh SS, Park E, Cho E, Ahn JH, Kim SG, Lee JS, Kwon YM, Lee I (2000) The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. Genes Dev 14 (18):2366-2376.

Lee JH, Park SH, Lee JS, Ahn JH (2007) A conserved role of SHORT VEGETATIVE PHASE (SVP) in controlling flowering time of Brassica plants. Biochim Biophys Acta Gene Struct Expr 1769 (7-8):455-461.

Lee JH, Ryu HS, Chung KS, Posé D, Kim S, Schmid M, Ahn JH (2013) Regulation of temperature-responsive flowering by MADS-box transcription factor repressors. Science 342 (6158):628-632.

Lee YS, An G (2015) Regulation of flowering time in rice. J Plant Biol 58 (6):353-360.

Lev-Yadun S, Gopher A, Abbo S (2000) The cradle of agriculture. Science 288 (5471):1602-1603.

Levy YY, Dean C (1998) The Transition to Flowering. The Plant Cell 10 (12):1973-1989.

Li P, Filiault D, Box MS, Kerdaffrec E, van Oosterhout C, Wilczek AM, Schmitt J, McMullan M, Bergelson J, Nordborg M, Dean C (2014) Multiple FLC haplotypes defined by independent cis-regulatory variation underpin life history diversity in Arabidopsis thaliana. Genes Dev 28 (15):1635-1640.

Li W, Wang H, Yu D (2016) Arabidopsis WRKY Transcription Factors WRKY12 and WRKY13 Oppositely Regulate Flowering under Short-Day Conditions. Mol Plant 9 (11):1492-1503.

Li Y, Ruperao P, Batley J, Edwards D, Davidson J, Hobson K, Sutton T (2017) Genome analysis identified novel candidate genes for ascochyta blight resistance in chickpea using whole genome re-sequencing data. Front Plant Sci 8.

Liang WH, Shang F, Lin QT, Lou C, Zhang J (2014) Tillering and panicle branching genes in rice. Gene 537 (1):1-5.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25 (11):1451-1452.

Lichtenzveig J, Bonfil DJ, Zhang H-B, Shtienberg D, Abbo S (2006) Mapping quantitative trait loci in chickpea associated with time to flowering and resistance to Didymella rabiei the causal agent of Ascochyta blight. Theoretical and Applied Genetics 113 (7):1357-1369.

Liew LC, Hecht V, Laurie RE, Knowles CL, Vander Schoor JK, Macknight RC, Weller JL (2009) DIE NEUTRALIS and LATE BLOOMER 1 contribute to regulation of the pea circadian clock. Plant Cell 21 (10):3198-3211.

Liew LC, Hecht V, Sussmilch FC, Weller JL (2014a) The pea photoperiod response gene STERILE NODES is an ortholog of LUX ARRHYTHMO. Plant Physiology 165 (2):648-657.

Liew LC, Singh MB, Bhalla PL (2014b) Unique and conserved features of floral evocation in legumes. J Integr Plant Biol.

Lifschitz E (2008) Multiple Regulatory Roles for SELF-PRUNING in the Shoot System of Tomato. Plant Physiology 148 (4):1737-1738.

Lifschitz E (2014) Florigen and anti-florigen – A systemic mechanism for coordinating growth and termination in flowering plants. Front Plant Sci 5 (SEP).

Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, Amsellem Z, Alvarez JP, Eshed Y (2006) The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. Proc Natl Acad Sci U S A 103 (16):6398-6403.

Lin C, Todo T (2005) The cryptochromes. Genome Biol 6 (5):220.

Liu B, Kanazawa A, Matsumura H, Takahashi R, Harada K, Abe J (2008) Genetic Redundancy in Soybean Photoresponses Associated With Duplication of the Phytochrome A Gene. Genetics 180 (2):995-1007.

Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, Xia Z, Nagamatsu A, Arai M, Yamada T, Kitamura K, Masuta C, Harada K, Abe J (2010) The soybean stem growth habit gene Dt1 is an ortholog of Arabidopsis TERMINAL FLOWER1. Plant Physiol 153 (1):198-210.

Liu L, Adrian J, Pankin A, Hu J, Dong X, von Korff M, Turck F (2014) Induced and natural variation of promoter length modulates the photoperiodic response of FLOWERING LOCUS T. Nat Commun 5.

Luan W, Chen H, Fu Y, Si H, Peng W, Song S, Liu W, Hu G, Sun Z, Xie D, Sun C (2009) The effect of the crosstalk between photoperiod and temperature on the heading-date in rice. Plos One 4 (6).

Mallikarjuna BP, Samineni S, Thudi M, Sajja SB, Khan AW, Patil A, Viswanatha KP, Varshney RK, Gaur PM (2017) Molecular Mapping of Flowering Time Major Genes and QTLs in Chickpea (Cicer arietinum L.). Front Plant Sci 8 (1140).

Mallikarjuna N, Coyne C, Cho S, Rynearson S, Rajesh PN, Jadhav D, Muehlbauer F (2011) Cicer. In: Kole C (ed) Wild Crop Relatives: Genomic and Breeding Resources. Springer Berlin Heidelberg, pp 63-82. doi:10.1007/978-3-642-14387-8_4

Mao Y, Sun J, Cao P, Zhang R, Fu Q, Chen S, Chen F, Jiang J (2016) Functional analysis of alternative splicing of the FLOWERING LOCUS T orthologous gene in Chrysanthemum morifolium. Horticulture Research 3:16058.

Martin-Trillo M, Grandio EG, Serra F, Marcel F, Rodriguez-Buey ML, Schmitz G, Theres K, Bendahmane A, Dopazo H, Cubas P (2011) Role of tomato BRANCHED1-like genes in the control of shoot branching. Plant J 67 (4):701-714.

Martinez-Garcia JF, Virgos-Soler A, Prat S (2002) Control of photoperiod-regulated tuberization in potato by the Arabidopsis flowering-time gene CONSTANS. Proc Natl Acad Sci U S A 99 (23):15211-15216.

Mathews S (2006) Phytochrome-mediated development in land plants: red light sensing evolves to meet the challenges of changing light environments. Mol Ecol 15 (12):3483-3503.

Mathieu J, Yant LJ, Mürdter F, Küttner F, Schmid M (2009) Repression of flowering by the miR172 target SMZ. PLoS Biology 7 (7).

Matsoukas IG (2015) Florigens and antiflorigens: a molecular genetic understanding. Essays in biochemistry 58:133-149.

Matsoukas IG, Massiah AJ, Thomas B (2012) Florigenic and Antflorigenic Signaling in Plants. Plant Cell Physiol 53 (11):1827-1842.

McClung CR (2014) Wheels within wheels: new transcriptional feedback loops in the Arabidopsis circadian clock. F1000prime reports 6:2.

McSteen P, Leyser O (2005) Shoot branching. Annu Rev Plant Biol 56:353-374.
Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T (2008) Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. Nat Genet 40.
Meng X, Muszynski MG, Danilevskaya ON (2011) The FT-Like ZCN8 Gene Functions as a Floral Activator and Is Involved in Photoperiod Sensitivity in Maize. The Plant Cell 23 (3):942-960.
Michaels SD, Amasino RM (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11 (5):949-956.
Michaels SD, Amasino RM (2001) Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of FRIGADA and autonomous pathway mutations but not responsiveness to vernalization. Plant Cell 13 (4):935-941.
Michaels SD, Bezerra IC, Amasino RM (2004) FRIGIDA-related genes are required for the winter-annual habit in Arabidopsis. Proc Natl Acad Sci U S A 101 (9):3281-3285.
Millan T, Clarke HJ, Siddique KHM, Buhariwalla HK, Gaur PM, Kumar J, Gil J, Kahl G, Winter P (2006) Chickpea molecular breeding: New tools and concepts. Euphytica 147 (1-2):81-103.
Millan T, Madrid E, Varshney K (2014) Genomic resources in chickpea. Legume perspectives (3):12-14.
Millan T, Winter P, Jüngling R, Gil J, Rubio J, Cho S, Cobos MJ, Iruela M, Rajesh PN, Tekeoglu M, Kahl G, Muehlbauer FJ (2010) A consensus genetic map of chickpea (Cicer arietinum L.) based on 10 mapping populations. Euphytica 175 (2):175-189.
Minakuchi K, Kameoka H, Yasuno N, Umehara M, Luo L, Kobayashi K, Hanada A, Ueno K, Asami T, Yamaguchi S, Kyozuka J (2010) FINE CULM1 (FC1) works downstream of strigolactones to inhibit the outgrowth of axillary buds in rice. Plant Cell Physiol 51 (7):1127-1135.
Mishra P, Panigrahi KC (2015) GIGANTEA – an emerging story. Front Plant Sci 6:8.
Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song HR, Carre IA, Coupland G (2002) LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis. Dev Cell 2 (5):629-641.
Mizoguchi T, Wright L, Fujiwara S, Cremer F, Lee K, Onouchi H, Mouradov A, Fowler S, Kamada H, Putterill J, Coupland G (2005) Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in Arabidopsis. Plant Cell 17 (8):2255-2270.
Mondal S, Singh RP, Crossa J, Huerta-Espino J, Sharma I, Chatrath R, Singh GP, Sohu VS, Mavi GS, Sukaru VSP, Kalappanavarg IK, Mishra VK, Hussain M, Gautam NR, Uddin J, Barma NCD, Hakim A, Joshi AK (2013) Earliness in wheat: A key to adaptation under terminal and continual high temperature stress in South Asia. Field Crops Research 151:19-26.
Monpara BA, Dhameliya HR (2013) Genetic behaviour of earliness related traits and seed yield in Chickpea (Cicer arietinum L.). Pakistan Journal of Biological Sciences 16 (18):955-959.
Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I (2003a) The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. Plant J 35 (5):613-623.
Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I (2003b) The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. Plant J 35 (5):613-623.
Moreno M-T, Cubero JI (1978) Variation in Cicer arietinum L. Euphytica 27 (2):465-485.
Morita MT (2010) Directional Gravity Sensing in Gravitropism. Annual Review of Plant Biology 61 (1):705-720.
Muehlbauer FJ, Singh KB (1987) Genetics of chickpea. In: Saxena MC, Singh KB (eds) The chickpea. C. A. B. International, Wallingford (K), pp 99-125
Nakamichi N (2011) Molecular Mechanisms Underlying the Arabidopsis Circadian Clock. Plant Cell Physiol 52 (10):1709-1718.
Nakamichi N, Kiba T, Henriques R, Mizuno T, Chua NH, Sakakibara H (2010) PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. Plant Cell 22 (3):594-605.
Nanda KK, Chinoy JJ (1960) Effect of vernalization and photoperiodic treatments on Cicer arietinum. 1. Phasic development in relation to its photo and thermic quanta. Indian Journal of Plant Physiology 3:32-44.

Nayak SN, Zhu H, Varghese N, Datta S, Choi HK, Horres R, Jüngling R, Singh J, Kishor PBK, Sivaramakrishnan S, Hoisington DA, Kahl G, Winter P, Cook DR, Varshney RK (2010) Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with Medicago truncatula genome. Theoretical and Applied Genetics 120 (7):1415-1441.

Nayyar H, Bains T, Kumar S (2005) Low temperature induced floral abortion in chickpea: Relationship to abscisic acid and cryoprotectants in reproductive organs. Environmental and Experimental Botany 53 (1):39-47.

Nelson MN, Książkiewicz M, Rychel S, Besharat N, Taylor CM, Wyrwa K, Jost R, Erskine W, Cowling WA, Berger JD, Batley J, Weller JL, Naganowska B, Wolko B (2017) The loss of vernalization requirement in narrow-leafed lupin is associated with a deletion in the promoter and de-repressed expression of a Flowering Locus T (FT) homologue. New Phytol 213 (1):220-232.

Nelson MN, Phan HTT, Ellwood SR, Moolhuijzen PM, Hane J, Williams A, O'Lone CE, Fosu-Nyarko J, Scobie M, Cakir M, Jones MGK, Bellgard M, Książkiewicz M, Wolko B, Barker SJ, Oliver RP, Cowling WA (2006) The first gene-based map of Lupinus angustifolius L.-location of domestication genes and conserved synteny with Medicago truncatula. Theoretical and Applied Genetics 113 (2):225-238.

Nelson R (1996) The Inheritance of a Branching Type in Soybean. Crop Science 36 (5):1150-1152.

Ngugi K, Collins JO, Muchira S (2013) Combining, earliness, short anthesis to silking interval and yield based selection indices under intermittent water stress to select for drought tolerant maize. Aust J Crop Sci 7 (13):2014-2020.

Nicholas KB, Nicholas HBJ (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author http://wwwpscedu/biomed/genedoc.

Ocampo B, Venora G, Errico A, Singh KB, Saccardo F (1992) Karyotype analysis in genus Cicer. Journal of Genetics and Breeding 46:229-240.

Ogawara T, Higashi K, Kamada H, Ezura H (2003) Ethylene advances the transition from vegetative growth to flowering in Arabidopsis thaliana. J Plant Physiol 160 (11):1335-1340.

Oliver SN, Deng W, Casao MC, Trevaskis B (2013) Low temperatures induce rapid changes in chromatin state and transcript levels of the cereal VERNALIZATION1 gene. J Exp Bot 64 (8):2413-2422.

Ongaro V, Leyser O (2008) Hormonal control of shoot branching. J Exp Bot 59 (1):67-74.

Ono N, Ishida K, Yamashino T, Nakanishi H, Sato S, Tabata S, Mizuno T (2010) Genomewide characterization of the light-responsive and clock-controlled output pathways in Lotus japonicus with special emphasis of its uniqueness. Plant Cell Physiol 51 (10):1800-1814.

Or E, Hovav R, Abbo S (1999) A major gene for flowering time in chickpea. Crop Science 39 (2):315-322.

Özdemir S, Karadavut U (2003) Comparison of the performance of autumn and spring sowing of chickpeas in a temperate region. Iliman şartlarda nohutun kişlik ekiminin yazlik ekime göre performansi 27 (6):345-352.

Pal BP, Murty GA (1941) Vernalization of Indian crops. Indian Journal of Genetics and Plant Breeding 1:61-85.

Palomino C, Fernández-Romero MD, Rubio J, Torres A, Moreno MT, Millán T (2009) Integration of new CAPS and dCAPS-RGA markers into a composite chickpea genetic map and their association with disease resistance. Theoretical and Applied Genetics 118 (4):671-682.

Pandey MK, Roorkiwal M, Singh VK, Ramalingam A, Kudapa H, Thudi M, Chitikineni A, Rathore A, Varshney RK (2016) Emerging Genomic Tools for Legume Breeding: Current Status and Future Prospects. Front Plant Sci 7 (455).

Patil PB, Vrinten PL, Scoles GJ, Slinkard AE (1995) Variation in the ribosomal RNA units of the genera Lens and Cicer. Euphytica 83 (1):33-42.

Pelaz S, Liljegren S, Roeder A, Ferrándiz C, Pinyopich A, Ostergaard L, Gremski K, Robles P, Ditta G, Kempin S, Yanofsky M (2003) The Role of MADS-Box Genes in the Control of Flower and Fruit Development in Arabidopsis. In: Vasil IK (ed) Plant Biotechnology 2002 and Beyond: Proceedings of the 10th IAPTC&B Congress June 23–28, 2002 Orlando, Florida, U.S.A. Springer Netherlands, Dordrecht, pp 20-27. doi:10.1007/978-94-017-2679-5_4

Penmetsa RV, Carrasquilla-Garcia N, Bergmann EM, Vance L, Castro B, Kassa MT, Sarma BK, Datta S, Farmer AD, Baek JM, Coyne CJ, Varshney RK, von Wettberg EJ, Cook DR (2016) Multiple post-domestication origins of kabuli chickpea through allelic variation in a diversification-associated transcription factor. New Phytol 211 (4):1440-1451.

Perilleux C, Pieltain A, Jacquemin G, Bouche F, Detry N, D'Aloia M, Thiry L, Aljochim P, Delansnay M, Mathieu AS, Lutts S, Tocquin P (2013) A root chicory MADS box sequence and the Arabidopsis flowering repressor FLC share common features that suggest conserved function in vernalization and de-vernalization responses. Plant J 75 (3):390-402.

Pfaff T, Kahl G (2003) Mapping of gene-specific markers on the genetic map of chickpea (Cicer arietinum L.). Mol Genet Genomics 269 (2):243-251.

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29 (9).

Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Mol Breed 7 (4):275-291.

Pierre JB, Huguet T, Barre P, Huyghe C, Julier B (2008) Detection of QTLs for flowering date in three mapping populations of the model legume species Medicago truncatula. Theor Appl Genet 117 (4):609-620.

Pin PA, Benlloch R, Bonnet D, Wremerth-Weich E, Kraft T, Gielen JJ, Nilsson O (2010) An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. Science 330 (6009):1397-1400.

Pin PA, Nilsson O (2012) The multifaceted roles of FLOWERING LOCUS T in plant development. Plant, Cell and Environment 35 (10):1742-1755.

Pin PA, Zhang W, Vogt SH, Dally N, Buttner B, Schulze-Buxloh G, Jelly NS, Chia TY, Mutasa-Gottgens ES, Dohm JC, Himmelbauer H, Weisshaar B, Kraus J, Gielen JJ, Lommel M, Weyens G, Wahl B, Schechert A, Nilsson O, Jung C, Kraft T, Muller AE (2012) The role of a pseudo-response regulator gene in life cycle adaptation and domestication of beet. Curr Biol 22 (12):1095-1101.

Ping J, Liu Y, Sun L, Zhao M, Li Y, She M, Sui Y, Lin F, Liu X, Tang Z, Nguyen H, Qiu L, Tian Z, Nelson RL, Clemente TE, Specht JE, Ma J (2014) Dt2 is a gain-of-function MADS-domain factor gene that specifies semideterminacy in soybean. Plant Cell 26 (7):2831-2842.

Pinhasi van-Oss R, Sherman A, Zhang HB, Vandemark G, Coyne C, Abbo S (2016) Vernalization response of domesticated × wild chickpea progeny is subject to strong genotype by environment interaction. Plant Breeding 135 (1):102-110.

Porri A, Torti S, Romera-Branchat M, Coupland G (2012) Spatially distinct regulatory roles for gibberellins in the promotion of flowering of arabidopsis under long photoperiods. Development (Cambridge) 139 (12):2198-2209.

Posé D, Verhage L, Ott F, Yant L, Mathieu J, Angenent GC, Immink RGH, Schmid M (2013) Temperature-dependent regulation of flowering by antagonistic FLM variants. Nature 503 (7476):414-417.

Poza-Carrion C, Aguilar-Martinez JA, Cubas P (2007) Role of TCP Gene BRANCHED1 in the Control of Shoot Branching in Arabidopsis. Plant Signal Behav 2 (6):551-552.

Proveniers MC, van Zanten M (2013) High temperature acclimation through PIF4 signaling. Trends Plant Sci 18 (2):59-64.

Pundir RPS, Reddy KN, Mengesha MH (1988) ICRISAT chickpea germplasm catalog : evaluation and analysis. ICRISAT, Patancheru

Purugganan MD, Suddith JI (1998) Molecular population genetics of the Arabidopsis CAULIFLOWER regulatory gene: Nonneutral evolution and naturally occurring variation in floral homeotic function. Proceedings of the National Academy of Sciences of the United States of America 95 (14):8130-8134.

Pushpavalli R, Krishnamurthy L, Thudi M, Gaur PM, Rao MV, Siddique KHM, Colmer TD, Turner NC, Varshney RK, Vadez V (2015) Two key genomic regions harbour QTLs for salinity tolerance in ICCV 2 × JG 11 derived chickpea (Cicer arietinum L.) recombinant inbred lines. Bmc Plant Biol 15 (1).

Putterill J, Robson F, Lee K, Simon R, Coupland G (1995a) The Constans Gene of Arabidopsis Promotes Flowering and Encodes a Protein Showing Similarities to Zinc-Finger Transcription Factors. Cell 80 (6):847-857.

Putterill J, Robson F, Lee K, Simon R, Coupland G (1995b) The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80 (6):847-857.

Putterill J, Varkonyi-Gasic E (2016) FT and florigen long-distance flowering control in plants. Curr Opin Plant Biol 33:77-82.

Putterill J, Zhang LL, Yeoh CC, Balcerowicz M, Jaudal M, Gasic EV (2013) FT genes and regulation of flowering in the legume Medicago truncatula. Funct Plant Biol 40 (12):1199-1207.

Qin Z, Wu J, Geng S, Feng N, Chen F, Kong X, Song G, Chen K, Li A, Mao L, Wu L (2017) Regulation of FT splicing by an endogenous cue in temperate grasses. Nature Communications 8:14320.

Raghu G (2001) The effect of earliness gene efl-1 on chickpea (Cicer arietinum L.) traits. Thesis,

Rajandran V (2016) Genetic control of flowering time in lentil (Doctoral dissertation). University of Tasmania,

Rajesh PN, Muehlbauer FJ (2008) Discovery and detection of single nucleotide polymorphism (SNP) in coding and genomic sequences in chickpea (Cicer arietinum L.). Euphytica 162 (2):291-300.

Rameau C, Bertheloot J, Leduc N, Andrieu B, Foucher F, Sakr S (2015) Multiple pathways regulate shoot branching. Front Plant Sci 5 (741).

Randoux M, Daviere JM, Jeauffre J, Thouroude T, Pierre S, Toualbia Y, Perrotte J, Reynoird JP, Jammes MJ, Hibrand-Saint Oyant L, Foucher F (2014) RoKSN, a floral repressor, forms protein complexes with RoFD and RoFT to regulate vegetative and reproductive development in rose. New Phytol 202 (1):161-173.

Ratcliffe OJ, Kumimoto RW, Wong BJ, Riechmann JL (2003) Analysis of the Arabidopsis MADS AFFECTING FLOWERING gene family: MAF2 prevents vernalization by short periods of cold. The Plant Cell 15 (5):1159-1169.

Ream TS, Woods DP, Amasino RM (2012) The molecular basis of vernalization in different plant groups. vol 77. doi:10.1101/sqb.2013.77.014449

Redden RJ, Berger JD (2007) History and origin of chickpea. In: Chickpea Breeding and Management. CABI Publishing, pp 1-13

Reeves PA, He Y, Schmitz RJ, Amasino RM, Panella LW, Richards CM (2007) Evolutionary Conservation of the FLOWERING LOCUS C-Mediated Vernalization Response: Evidence From the Sugar Beet (Beta vulgaris). Genetics 176 (1):295-307.

Rehman AU, Malhotra RS, Bett K, Tar'an B, Bueckert R, Warkentin TD (2011) Mapping QTL associated with traits affecting grain yield in chickpea (Cicer arietinum L.) under terminal drought stress. Crop Science 51 (2):450-463.

Reinhardt D, Kuhlemeier C (2002) Plant architecture. Embo Rep 3 (9):846-851.

Ridge S, Deokar A, Lee R, Daba K, Macknight RC, Weller JL, Tar, an B (2017) The chickpea Early Flowering 1 (Efl1) locus is an ortholog of Arabidopsis ELF3. Plant Physiology.

Ridge S, Sussmilch FC, Hecht VF, Vander Schoor JK, Lee R, Aubert G, Burstin J, Macknight RC, Weller JL (2016) Identification of LATE BLOOMER2 as a CYCLING DOF FACTOR Homolog Reveals Conserved and Divergent Features of the Flowering Response to Photoperiod in Pea. The Plant Cell.

Roberts EH, Hadley P, Summerfield RJ (1985) Effects of temperature and photoperiod on flowering in Chickpeas (Cicer arietinum L.). Annals of Botany 55 (6):881-892.

Roberts EH, Summerfield RJ, Minchin FR, Hadley P (1980) Phenology of Chickpeas (*Cicer arietinum L.*) in contrasting aerial environments. Experimental Agriculture 16:343-360.

Robins JG, Bauchan GR, Brummer EC (2007) Genetic mapping forage yield, plant height, and regrowth at multiple harvests in tetraploid Alfalfa (Medicago sativa L.). Crop Science 47 (1):11-18.

Robson F, Costa MM, Hepworth SR, Vizir I, Pineiro M, Reeves PH, Putterill J, Coupland G (2001) Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. Plant J 28 (6):619-631.

Roychoudhry S, Del Bianco M, Kieffer M, Kepinski S (2013) Auxin controls gravitropic setpoint angle in higher plant lateral branches. Curr Biol 23 (15):1497-1504.

Roychoudhry S, Kepinski S (2015) Shoot and root branch growth angle control-the wonderfulness of lateralness. Curr Opin Plant Biol 23:124-131.

Rubenach AJ, Hecht V, Vander Schoor JK, Liew LC, Aubert G, Burstin J, Weller JL (2017) EARLY FLOWERING3 Redundancy Fine-Tunes Photoperiod Sensitivity. Plant Physiol 173 (4):2253-2264.

Rubio J, Flores F, Moreno MT, Cubero JI, Gil J (2004) Effects of the erect/bushy habit, single/double pod and late/early flowering genes on yield and seed size and their stability in chickpea. Field Crops Research 90 (2-3):255-262.

Ruelens P, de Maagd RA, Proost S, Theissen G, Geuten K, Kaufmann K (2013) FLOWERING LOCUS C in monocots and the tandem origin of angiosperm-specific MADS-box genes. Nat Commun 4:2280.

Sadeghipour O, Aghaei P (2012) Comparison of autumn and spring sowing on performance of chickpea (Cicer arietinum L.) varieties. International Journal of Biosciences 2 (6):49-58.

Saha GC, Vandemark GJ (2012) Evaluation of expression stability of candidate references genes among green and yellow pea cultivars (Pisum sativum L.) subjected to abiotic and biotic stress. American Journal of Plant Sciences 3 (2):235.

Saha GC, Vandemark GJ (2013) Stability of Expression of Reference Genes Among Different Lentil (Lens culinaris) Genotypes Subjected to Cold Stress, White Mold Disease, and Aphanomyces Root Rot. Plant Mol Biol Report 31 (5):1109-1115.

Samineni S, Kamatam S, Thudi M, Varshney RK, Gaur PM (2016) Vernalization response in chickpea is controlled by a major QTL. Euphytica 207 (2):453-461.

Saxena MC, Siddique M (1980) Response of some diverse kabuli chickpea genotypes to vernalization. International Chickpea Newsletter (No.2):7-8.

Saxena MS, Bajaj D, Das S, Kujur A, Kumar V, Singh M, Bansal KC, Tyagi AK, Parida SK (2014) An integrated genomic approach for rapid delineation of candidate genes regulating agro-morphological traits in chickpea. DNA Research 21 (6):695-710.

Schiessl S, Iniguez-Luy F, Qian W, Snowdon RJ (2015) Diverse regulatory factors associate with flowering time and yield responses in winter-type Brassica napus. BMC Genomics 16 (1).

Schlappi MR (2006) FRIGIDA LIKE 2 is a functional allele in Landsberg erecta and compensates for a nonsense allele of FRIGIDA LIKE 1. Plant Physiol 142 (4):1728-1738.

Schmitz RJ, Amasino RM (2007) Vernalization: a model for investigating epigenetics and eukaryotic gene regulation in plants. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression 1769 (5):269-275.

Schomburg FM, Bizzell CM, Lee DJ, Zeevaart JA, Amasino RM (2003) Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants. Plant Cell 15 (1):151-163.

Schwartz C, Balasubramanian S, Warthmann N, Michael TP, Lempe J, Sureshkumar S, Kobayashi Y, Maloof JN, Borevitz JO, Chory J, Weigel D (2009) Cis-regulatory Changes at FLOWERING LOCUS T Mediate Natural Variation in Flowering Responses of Arabidopsis thaliana. Genetics 183 (2):723-732.

Searle I, Coupland G (2004) Induction of flowering by seasonal changes in photoperiod. The EMBO Journal 23 (6):1217-1222.

Searle I, He Y, Turck F, Vincent C, Fornara F, Kröber S, Amasino RA, Coupland G (2006) The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. Gene Dev 20 (7):898-912.

Serrano G, Herrera-Palau R, Romero JM, Serrano A, Coupland G, Valverde F (2009) Chlamydomonas CONSTANS and the evolution of plant photoperiodic signaling. Curr Biol 19 (5):359-368.

Sethi SC, Byth DE, Gowda CLL, Green JM (1981) Photoperiodic response and accelerated generation turnover in chickpea. Field Crops Research 4:215-225.

Shalit A, Rozman A, Goldshmidt A, Alvarez JP, Bowman JL, Eshed Y, Lifschitz E (2009) The flowering hormone florigen functions as a general systemic regulator of growth and termination. Proceedings of the National Academy of Sciences 106 (20):8392-8397.

Sharma S, Upadhyaya HD (2015) Vernalization and photoperiod response in annual wild cicer species and cultivated chickpea. Crop Science 55 (5):2393-2400.

Sharma S, Upadhyaya HD, Roorkiwal M, Varshney RK, Gowda CLL (2013) 4 - Chickpea. In: Genetic and Genomic Resources of Grain Legume Improvement. Elsevier, Oxford, pp 81-111. doi:https://doi.org/10.1016/B978-0-12-397935-3.00004-9

Shi J, Li R, Qiu D, Jiang C, Long Y, Morgan C (2009) Unraveling the Complex Trait of Crop Yield With Quantitative Trait Loci Mapping in Brassica napus. Genetics 182.

Shim JS, Imaizumi T (2015) Circadian clock and photoperiodic response in arabidopsis: From seasonal flowering to redox homeostasis. Biochemistry 54 (2):157-170.

Shrestha R, Gómez-Ariza J, Brambilla V, Fornara F (2014) Molecular control of seasonal flowering in rice, arabidopsis and temperate cereals. Annals of Botany 114 (7):1445-1458.

Shu K, Chen Q, Wu Y, Liu R, Zhang H, Wang S, Tang S, Yang W, Xie Q (2016) ABSCISIC ACID-INSENSITIVE 4 negatively regulates flowering through directly promoting Arabidopsis FLOWERING LOCUS C transcription. Journal of Experimental Botany 67 (1):195-205.

Siddique KHM, Brinsmead RB, Knight R, Knights EJ, Paull JG, Rose IA (2000) Adaptation of chickpea (Cicer arietinum L.) and faba bean (Vicia faba L.) to Australia. In: Knight R (ed) Linking Research and Marketing Opportunities for Pulses in the 21st Century: Proceedings of the Third International Food Legumes Research Conference. Springer Netherlands, Dordrecht, pp 289-303. doi:10.1007/978-94-011-4385-1_26

Siddique KHM, Krishnamurthy L (2014) Chickpea production technology. Legume perspectives (3):29-32.

Siddique KHM, Sedgley RH (1985) The effect of reduced branching on yield and water use of chickpea (Cicer arietinum L.) in a Mediterranean type environment. Field Crops Research 12 (C):251-269.

Simon CJ, Muehlbauer FJ (1997) Construction of a chickpea linkage map and its comparison with maps of pea and lentil. Journal of Heredity 88 (2):115-119.

Singh D, Shyam R (1959) Genetics of two new mutants in *Cicer arietinum*. Indian Journal of Genetics and Plant Breeding (The) 19:73-82.

Singh KB (1997) Chickpea (Cicer arietinum L). Field Crops Research 53 (1-3):161-170.

Singh KB, Malhotra RS, Halila MH, Knights EJ, Verma MM (1993) Current status and future strategy in breeding chickpea for resistance to biotic and abiotic stresses. Euphytica 73 (1-2):137-149.

Singh KB, Malhotra RS, Saxena MC, Bejiga G (1997) Superiority of winter sowing over traditional spring sowing of chickpea in the mediterranean region. Agronomy Journal 89 (1):112-118.

Singh KB, Ocampo B (1993) Interspecific hybridization in annual *Cicer* species. Journal of Genetics and Breeding 47:199-204.

Singh KB, Ocampo B (1997) Exploitation of wild Cicer species for yield improvement in chickpea. Theoretical and Applied Genetics 95 (3):418-423.

Singh KB, Reddy MV (1996) Improving chickpea yield by incorporating resistance to ascochyta blight. Theor Appl Genet 92 (5):509-515.

Singh S, Gumber RK, Joshi N, Singh K (2005) Introgression from wild Cicer reticulatum to cultivated chickpea for productivity and disease resistance. Plant Breeding 124 (5):477-480.

Singh VK, Garg R, Jain M (2013) A global view of transcriptome dynamics during flower development in chickpea by deep sequencing. Plant Biotechnol J 11 (6):691-701.

Soltani A, Hammer GL, Torabi B, Robertson MJ, Zeinali E (2006) Modeling chickpea growth and development: Phenological development. Field Crops Research 99 (1):1-13.

Soltani A, Torabi B, Zeinali E, Sarparast R (2004) Response of chickpea to photoperiod as a qualitative long-day plant. Asian J Plant Sci 3 (6):705-708.

Somers DE, Devlin PF, Kay SA (1998) Phytochromes and cryptochromes in the entrainment of the Arabidopsis circadian clock. Science 282 (5393):1488-1490.

Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T (2015) Photoperiodic flowering: time measurement mechanisms in leaves. Annu Rev Plant Biol 66:441-464.

Sorefan K, Booker J, Haurogne K, Goussot M, Bainbridge K, Foo E, Chatfield S, Ward S, Beveridge C, Rameau C, Leyser O (2003) MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea. Genes Dev 17 (12):1469-1474.

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. Plant J 3 (5):739-744.
Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z, Toiber D, Thanaraj TA, Soreq H (2005) Function of alternative splicing. Gene 344 (Supplement C):1-20.
Strasser B, Alvarez MJ, Califano A, Cerdán PD (2009) A complementary role for ELF3 and TFL1 in the regulation of flowering time by ambient temperature. Plant J 58 (4):629-640.
Strayer C, Oyama T, Schultz TF, Raman R, Somers DE, Mas P, Panda S, Kreps JA, Kay SA (2000) Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. Science 289 (5480):768-771.
Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. Nat Genet 43 (11):1160-1163.
Suárez-López P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature 410 (6832):1116-1120.
Subbarao GV, Johansen C, Slinkard AE, Nageswara Rao RC, Saxena NP, Chauhan YS (1995) Strategies For Improving Drought Resistance In Grain Legumes. Crit Rev Plant Sci 14 (6):469-523.
Sudupak MA (2004) Inter and intra-species Inter Simple Sequence Repeat (ISSR) variations in the genus Cicer. Euphytica 135 (2):229-238.
Sudupak MA, Akkaya MS, Kence A (2002) Analysis of genetic relationships among perennial and annual Cicer species growing in Turkey using RAPD markers. Theoretical and Applied Genetics 105 (8):1220-1228.
Summerfield RJ, Ellis RH, Roberts EH (1989) Vernalization in Chickpea (Cicer arietinum); Fact or Artefact? Annals of Botany 64 (5):599-603.
Summerfield RJ, Minchin FR, Roberts EH, Hadley P (1981) Adaptation to contrasting aerial environments in chickpea (Cicer arietinum L.). Tropical Agriculture 58:97-113.
Summerfield RJ, Roberts EH (1988) Photo-thermal regulation of flowering in pea, lentil, faba bean and chickpea. In: Summerfield RJ (ed) World crops: Cool season food legumes, vol 5. Current Plant Science and Biotechnology in Agriculture. Springer Netherlands, pp 911-922. doi:10.1007/978-94-009-2764-3_72
Sun H, Jia Z, Cao D, Jiang B, Wu C, Hou W, Liu Y, Fei Z, Zhao D, Han T (2011) GmFT2a, a Soybean Homolog of FLOWERING LOCUS T, Is Involved in Flowering Transition and Maintenance. Plos One 6 (12):e29238.
Sung S, Amasino RM (2004) Vernalization and epigenetics: how plants remember winter. Curr Opin Plant Biol 7 (1):4-10.
Sung S, Amasino RM (2006) Molecular genetic studies of the memory of winter. Journal of Experimental Botany 57 (13):3369-3377.
Sung S, Schmitz RJ, Amasino RM (2006) A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis. Genes Dev 20 (23):3244-3248.
Sussmilch FC, Berbel A, Hecht V, Vander Schoor JK, Ferrándiz C, Madueño F, Weller JL (2015) Pea VEGETATIVE2 is an FD homolog that is essential for flowering and compound inflorescence development. Plant Cell 27 (4):1046-1060.
Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature 462 (7274):799-802.
Swofford DL (2001) Paup*: Phylogenetic analysis using parsimony (and other methods) 4.0. B5.
Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchi-Tanaka M, Ashikari M, Matsuoka M, Ueguchi C (2003) The OsTB1 gene negatively regulates lateral branching in rice. Plant J 33 (3):513-520.
Takeshima R, Hayashi T, Zhu J, Zhao C, Xu M, Yamaguchi N, Sayama T, Ishimoto M, Kong L, Shi X, Liu B, Tian Z, Yamada T, Kong F, Abe J (2016) A soybean quantitative trait locus that promotes flowering under long days is identified as FT5a, a FLOWERING LOCUS T ortholog. Journal of Experimental Botany 67 (17):5247-5258.
Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K (2007) Hd3a protein is a mobile flowering signal in rice. Science 316 (5827):1033-1036.
Tan J, Jin M, Wang J, Wu F, Sheng P, Cheng Z, Wang J, Zheng X, Chen L, Wang M, Zhu S, Guo X, Zhang X, Liu X, Wang C, Wang H, Wu C, Wan J (2016) OsCOL10 , a CONSTANS-Like

Gene, Functions as a Flowering Time Repressor Downstream of Ghd7 in Rice. Plant Cell Physiol 57 (4):798-812.

Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzbittel L, Childs KL, Yandell M, Gundlach H, Mayer KF, Schwartz DC, Town CD (2014) An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC Genomics 15 (1):312.

Taoka K, Ohki I, Tsuji H, Furuita K, Hayashi K, Yanase T, Yamaguchi M, Nakashima C, Purwestri YA, Tamaki S, Ogaki Y, Shimada C, Nakagawa A, Kojima C, Shimamoto K (2011) 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. Nature 476 (7360):332-335.

Teichmann T, Muhr M (2015) Shaping plant architecture. Front Plant Sci 6:233.

Tekeoglu M, Rajesh N, Muehlbauer J (2002) Integration of sequence tagged microsatellite sites to the chickpea genetic map. Theor Appl Genet 105 (6-7):847-854.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22 (22):4673-4680.

Thudi M, Bohra A, Nayak SN, Varghese N, Shah TM, Penmetsa RV, Thirunavukkarasu N, Gudipati S, Gaur PM, Kulwal PL, Upadhyaya HD, KaviKishor PB, Winter P, Kahl G, Town CD, Kilian A, Cook DR, Varshney RK (2011) Novel SSR Markers from BAC-End Sequences, DArT Arrays and a Comprehensive Genetic Map with 1,291 Marker Loci for Chickpea (Cicer arietinum L.). Plos One 6 (11).

Thudi M, Khan AW, Kumar V, Gaur PM, Katta K, Garg V, Roorkiwal M, Samineni S, Varshney RK (2016) Whole genome re-sequencing reveals genome-wide variations among parental lines of 16 mapping populations in chickpea (Cicer arietinum L.). Bmc Plant Biol 16 (1).

Thudi M, Upadhyaya HD, Rathore A, Gaur PM, Krishnamurthy L, Roorkiwal M, Nayak SN, Chaturvedi SK, Basu PS, Gangarao NVPR, Fikre A, Kimurto P, Sharma PC, Sheshashayee MS, Tobita S, Kashiwagi J, Ito O, Killian A, Varshney RK (2014) Genetic Dissection of Drought and Heat Tolerance in Chickpea through Genome-Wide and Candidate Gene-Based Association Mapping Approaches. Plos One 9 (5):e96758.

Tian Z, Wang X, Lee R, Li Y, Specht JE, Nelson RL, McClean PE, Qiu L, Ma J (2010) Artificial selection for determinate growth habit in soybean. Proceedings of the National Academy of Sciences 107 (19):8563-8568.

Tiwari SB, Shen Y, Chang H-C, Hou Y, Harris A, Ma SF, McPartland M, Hymus GJ, Adam L, Marion C, Belachew A, Repetti PP, Reuber TL, Ratcliffe OJ (2010) The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. New Phytol 187 (1):57-66.

Tsuji H, Tachibana C, Tamaki S, Taoka K, Kyozuka J, Shimamoto K (2015) Hd3a promotes lateral branching in rice. Plant J 82 (2):256-266.

Turck F, Fornara F, Coupland G (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. Annu Rev Plant Biol 59:573-594.

Turner NC, Wright GC, Siddique KHM (2001) Adaptation of grain legumes (pulses) to water-limited environments. Advances in Agronomy, vol 71.

Upadhyaya HD, Bajaj D, Das S, Saxena MS, Badoni S, Kumar V, Tripathi S, Gowda CL, Sharma S, Tyagi AK, Parida SK (2015) A genome-scale integrated approach aids in genetic dissection of complex flowering time trait in chickpea. Plant Mol Biol 89 (4-5):403-420.

Upadhyaya HD, Bajaj D, Narnoliya L, Das S, Kumar V, Gowda CL, Sharma S, Tyagi AK, Parida SK (2016) Genome-Wide Scans for Delineation of Candidate Genes Regulating Seed-Protein Content in Chickpea. Front Plant Sci 7:302.

Upadhyaya HD, Bajaj D, Srivastava R, Daware A, Basu U, Tripathi S, Bharadwaj C, Tyagi AK, Parida SK (2017) Genetic dissection of plant growth habit in chickpea. Functional & integrative genomics.

Upadhyaya HD, Furman BJ, Dwivedi SL, Udupa SM, Gowda CLL, Baum M, Crouch JH, Buhariwalla HK, Singh S (2006) Development of a composite collection for mining germplasm possessing allelic variation for beneficial traits in chickpea. Plant Genetic Resources 4 (1):13-19.

Upadhyaya HD, Thudi M, Dronavalli N, Gujaria N, Singh S, Sharma S, Varshney RK (2011) Genomic tools and germplasm diversity for chickpea improvement. Plant Genetic Resources 9 (1):45-58.

Ustyantsev K, Novikova O, Blinov A, Smyshlyaev G (2015) Convergent Evolution of Ribonuclease H in LTR Retrotransposons and Retroviruses. Molecular Biology and Evolution 32 (5):1197-1207.

Vadez V, Krishnamurthy L, Thudi M, Anuradha C, Colmer TD, Turner NC, Siddique KHM, Gaur PM, Varshney RK (2012) Assessment of ICCV 2 × JG 62 chickpea progenies shows sensitivity of reproduction to salt stress and reveals QTL for seed yield and yield components. Mol Breed 30 (1):9-21.

Valimohammadi F, Tajbakhsh M, Saeid A (2007) Comparison winter and spring sowing dates and effect of plant density on yield, yield components and some quality, morphological traits of chickpea (Cicer arietinum L.) under environmental condition of Urmia, Iran. J Agron 6 (4):571-575.

Valverde F (2011) CONSTANS and the evolutionary origin of photoperiodic timing of flowering. Journal of experimental botany 62 (8):2453-2463.

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G (2004) Photoreceptor Regulation of CONSTANS Protein in Photoperiodic Flowering. Science 303 (5660):1003-1006.

Van Der Maesen LJG (1972) Cicer L., a monograph of the genus, with special reference to the chickpea (Cicer arietinum L.), its ecology and cultivation.

Van Der Maesen LJG Origin, history and taxonomy of Chickpea. In: Saxena MC, Singh KB (eds) The chickpea, Wallingford, 1987. C.A.B. International, pp 11-34

Van Der Maesen LJG, Maxted N, Javadi F, Coles S, Davies AMR (2007) Taxonomy of the genus Cicer revisited. In: Chickpea Breeding and Management. CABI Publishing, pp 14-46

Van Ooijen JW (2006) JoinMap 4: Software for the calculation of genetic linkage maps in experimental populations (Kyazma B.V., Wageningen).

Van Ooijen JW, Maliepaard (1996) MapQTL version 3.0: Software for the calculation of QTL positions on genetic maps. Plant Genome IV Abstracts

Varkonyi-Gasic E, Moss SMA, Voogd C, Wang TC, Putterill J, Hellens RP (2013) Homologs of FT, CEN and FD respond to developmental and environmental signals affecting growth and flowering in the perennial vine kiwifruit. New Phytol 198 (3):732-746.

Varshney R, Thudi M, Nayak S, Gaur P, Kashiwagi J, Krishnamurthy L, Jaganathan D, Koppolu J, Bohra A, Tripathi S, Rathore A, Jukanti A, Jayalakshmi V, Vemula A, Singh SJ, Yasin M, Sheshshayee MS, Viswanatha KP (2014a) Genetic dissection of drought tolerance in chickpea (Cicer arietinum L.). Theoretical and Applied Genetics 127 (2):445-462.

Varshney RK, Close TJ, Singh NK, Hoisington DA, Cook DR (2009a) Orphan legume crops enter the genomics era! Current Opinion in Plant Biology 12 (2):202-210.

Varshney RK, Hiremath PJ, Lekha P, Kashiwagi J, Balaji J, Deokar AA, Vadez V, Xiao Y, Srinivasan R, Gaur PM, Siddique KH, Town CD, Hoisington DA (2009b) A comprehensive resource of drought- and salinity- responsive ESTs for gene discovery and marker development in chickpea (Cicer arietinum L.). BMC Genomics 10 (523):1471-2164.

Varshney RK, Mohan SM, Gaur PM, Gangarao NVPR, Pandey MK, Bohra A, Sawargaonkar SL, Chitikineni A, Kimurto PK, Janila P, Saxena KB, Fikre A, Sharma M, Rathore A, Pratap A, Tripathi S, Datta S, Chaturvedi SK, Mallikarjuna N, Anuradha G, Babbar A, Choudhary AK, Mhase MB, Bharadwaj C, Mannur DM, Harer PN, Guo B, Liang X, Nadarajan N, Gowda CLL (2013a) Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. Biotechnol Adv 31 (8):1120-1134.

Varshney RK, Nayak SN, May GD, Jackson SA (2009c) Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends in biotechnology 27 (9):522-530.

Varshney RK, Roorkiwal M, Nguyen HT (2013b) Legume Genomics: From Genomic Resources to Molecular Breeding. Plant Gen 6 (3):-.

Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, Cannon S, Baek J, Rosen BD, Tar'an B, Millan T, Zhang X, Ramsay LD, Iwata A, Wang Y, Nelson W, Farmer AD, Gaur PM, Soderlund C, Penmetsa RV, Xu C, Bharti AK, He W, Winter P, Zhao S, Hane JK,

Carrasquilla-Garcia N, Condie JA, Padhyaya HD, Luo MC, Thudi M, Gowda CLL, Singh NP, Lichtenzveig J, Gali KK, Rubio J, Nadarajan N, Dolezel J, Bansal KC, Xu X, Edwards D, Zhang G, Kahl G, Gil J, Ingh KB, Datta SK, Jackson SA, Wang J, Cook DR (2013c) Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nat Biotechnol 31 (3):240-246.

Varshney RK, Thudi M, Nayak SN, Gaur PM, Kashiwagi J, Krishnamurthy L, Jaganathan D, Koppolu J, Bohra A, Tripathi S, Rathore A, Jukanti AK, Jayalakshmi V, Vemula A, Singh SJ, Yasin M, Sheshshayee MS, Viswanatha KP (2014b) Genetic dissection of drought tolerance in chickpea (Cicer arietinum L.). Theoretical and Applied Genetics 127 (2):445-462.

Verma P, Kaur H, Petla BP, Rao V, Saxena SC, Majee M (2013) PROTEIN L-ISOASPARTYL METHYLTRANSFERASE2 is differentially expressed in chickpea and enhances seed vigor and longevity by reducing abnormal isoaspartyl accumulation predominantly in seed nuclear proteins. Plant Physiol 161 (3):1141-1157.

Vogt SH, Weyens G, Lefèbvre M, Bork B, Schechert A, Müller AE (2014) The FLC-like gene BvFL1 is not a major regulator of vernalization response in biennial beets. Front Plant Sci 5:146.

von Zitzewitz J, Szűcs P, Dubcovsky J, Yan L, Francia E, Pecchioni N, Casas A, Chen TH, Hayes PM, Skinner JS (2005) Molecular and structural characterization of barley vernalization genes. Plant Mol Biol 59 (3):449-467.

Voogd C, Brian LA, Wang T, Allan AC, Varkonyi-Gasic E (2017) Three FT and multiple CEN and BFT genes regulate maturity, flowering, and vegetative phenology in kiwifruit. Journal of Experimental Botany 68 (7):1539-1553.

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93 (1):77-78.

Wang BB, Brendel V (2006) Genomewide comparative analysis of alternative splicing in plants. Proc Natl Acad Sci U S A 103 (18):7175-7180.

Wang H, Huang J, Lai Z, Xue Y (2002) F-box proteins in flowering plants. Chinese Science Bulletin 47 (18):1497-1501.

Wang Y, Li J (2008) Molecular basis of plant architecture. Annu Rev Plant Biol 59:253-279.

Wang Y, Li L, Ye T, Lu Y, Chen X, Wu Y (2013) The inhibitory effect of ABA on floral transition is mediated by ABI5 in Arabidopsis. Journal of Experimental Botany 64 (2):675-684.

Wang Y, Wu JF, Nakamichi N, Sakakibara H, Nam HG, Wu SH (2011) LIGHT-REGULATED WD1 and PSEUDO-RESPONSE REGULATOR9 form a positive feedback regulatory loop in the Arabidopsis circadian clock. Plant Cell 23 (2):486-498.

Wang Z, Zhou Z, Liu Y, Liu T, Li Q, Ji Y, Li C, Fang C, Wang M, Wu M, Shen Y, Tang T, Ma J, Tian Z (2015) Functional evolution of phosphatidylethanolamine binding proteins in soybean and Arabidopsis. Plant Cell 27 (2):323-336.

Ward SP, Leyser O (2004) Shoot branching. Current Opinion in Plant Biology 7 (1):73-78.

Wasilewska A, Vlad F, Sirichandra C, Redko Y, Jammes F, Valon C, Frey NFd, Leung J (2008) An Update on Abscisic Acid Signaling in Plants and More …. Mol Plant 1 (2):198-217.

Watanabe S, Hideshima R, Xia Z, Tsubokura Y, Sato S, Nakamoto Y, Yamanaka N, Takahashi R, Ishimoto M, Anai T, Tabata S, Harada K (2009) Map-based cloning of the gene associated with the soybean maturity locus E3. Genetics 182 (4):1251-1262.

Watanabe S, Xia Z, Hideshima R, Tsubokura Y, Sato S, Yamanaka N, Takahashi R, Anai T, Tabata S, Kitamura K, Harada K (2011) A map-based cloning strategy employing a residual heterozygous line reveals that the GIGANTEA gene is involved in soybean maturity and flowering. Genetics 188 (2):395-407.

Weller JL, Batge SL, Smith JJ, Kerckhoffs LHJ, Sineshchekov VA, Murfet IC, Reid JB (2004) A Dominant Mutation in the Pea PHYA Gene Confers Enhanced Responses to Light and Impairs the Light-Dependent Degradation of Phytochrome A. Plant Physiology 135 (4):2186-2195.

Weller JL, Liew LC, Hecht VFG, Rajandran V, Laurie RE, Ridge S, Wenden B, Schoor JKV, Jaminon O, Blassiau C, Dalmais M, Rameau C, Bendahmane A, Macknight RC, Lejeune-Hénaut I (2012) A conserved molecular basis for photoperiod adaptation in two temperate legumes. Proceedings of the National Academy of Sciences of the United States of America 109 (51):21158-21163.

Weller JL, Murfet IC, Reid JB (1997a) Pea Mutants with Reduced Sensitivity to Far-Red Light Define an Important Role for Phytochrome A in Day-Length Detection. Plant Physiol 114 (4):1225-1236.

Weller JL, Ortega R (2015) Genetic control of flowering time in legumes. Front Plant Sci 6 (207).

Weller JL, Reid JB, Taylor SA, Murfet IC (1997b) The genetic control of flowering in pea. Trends in Plant Science 2 (11):412-418.

Weng L, Bai X, Zhao F, Li R, Xiao H (2016) Manipulation of flowering time and branching by overexpression of the tomato transcription factor SlZFP2. Plant Biotechnol J:n/a-n/a.

Wickland DP, Hanzawa Y (2015) The FLOWERING LOCUS T / TERMINAL FLOWER 1 Gene Family: Functional Evolution and MolecularMechanisms. Mol Plant 8 (7):983-997.

Wigge PA (2013) Ambient temperature signalling in plants. Curr Opin Plant Biol 16.

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D (2005) Integration of spatial and temporal information during floral induction in Arabidopsis. Science 309 (5737):1056-1059.

Wilson RN, Heckman JW, Somerville CR (1992) Gibberellin is required for flowering in Arabidopsis thaliana under short days. Plant Physiology 100 (1):403-408.

Winter P, Benko-Iseppon AM, Hüttel B, Ratnaparkhe M, Tullu A, Sonnante G, Pfaff T, Tekeoglu M, Santra D, Sant VJ, Rajesh PN, Kahl G, Muehlbauer FJ (2000) A linkage map of the chickpea (Cicer arietinum L.) genome based on recombinant inbred lines from a C. arietinum x C. reticulatum cross: Localization of resistance genes for fusarium wilt races 4 and 5. Theoretical and Applied Genetics 101 (7):1155-1163.

Winter P, Pfaff T, Udupa SM, Hüttel B, Sharma PC, Sahi S, Arreguin-Espinoza R, Weigand F, Muehlbauer FJ, Kahl G (1999) Characterization and mapping of sequence-tagged microsatellite sites in the chickpea (Cicer arietinum L.) genome. Molecular and General Genetics 262 (1):90-101.

Wong ACS, Hecht VFG, Picard K, Diwadkar P, Laurie RE, Wen J, Mysore K, Macknight RC, Weller JL (2014) Isolation and functional analysis of CONSTANS-LIKE genes suggests that a central role for CONSTANS in flowering time control is not evolutionarily conserved in Medicago truncatula. Front Plant Sci 5.

Wood JA, Grusak MA (2007) Nutritional value of chickpea. In: Chickpea Breeding and Management. CABI Publishing, pp 101-142

Wu F, Price BW, Haider W, Seufferheld G, Nelson R, Hanzawa Y (2014) Functional and Evolutionary Characterization of the CONSTANS Gene Family in Short-Day Photoperiodic Flowering in Soybean. Plos One 9 (1):e85754.

Wu G, Poethig RS (2006) Temporal regulation of shoot development in Arabidopsis thaliana by miRr156 and its target SPL3. Development 133 (18):3539-3547.

Wu HJ, Wang ZM, Wang M, Wang XJ (2013a) Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. Plant Physiology 161 (4):1875-1884.

Wu JF, Tsai HL, Joanito I, Wu YC, Chang CW, Li YH, Wang Y, Hong JC, Chu JW, Hsu CP, Wu SH (2016) LWD-TCP complex activates the morning gene CCA1 in Arabidopsis. Nat Commun 7:13181.

Wu JF, Wang Y, Wu SH (2008) Two new clock proteins, LWD1 and LWD2, regulate Arabidopsis photoperiodic flowering. Plant Physiol 148 (2):948-959.

Wu KL, Guo ZJ, Wang HH, Li J (2005) The WRKY family of transcription factors in rice and Arabidopsis and their origins. DNA research : an international journal for rapid publication of reports on genes and genomes 12 (1):9-26.

Wu L, Liu D, Wu J, Zhang R, Qin Z, Liu D, Li A, Fu D, Zhai W, Mao L (2013b) Regulation of *FLOWERING LOCUS T* by a MicroRNA in *Brachypodium distachyon*. The Plant Cell 25 (11):4363-4377.

Xu M, Xu Z, Liu B, Kong F, Tsubokura Y, Watanabe S, Xia Z, Harada K, Kanazawa A, Yamada T, Abe J (2013) Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. Bmc Plant Biol 13:91.

Yadav SS, Redden RJ, Chen W, Sharma B Chickpea breeding and management. Book.

Yamaguchi A, Kobayashi Y, Goto K, Abe M, Araki T (2005) TWIN SISTER OF FT (TSF) acts as a floral pathway integrator redundantly with FT. Plant Cell Physiol 46 (8):1175-1189.

Yamashino T, Yamawaki S, Hagui E, Ueoka-Nakanishi H, Nakamichi N, Ito S, Mizuno T (2013) Clock-controlled and FLOWERING LOCUS T (FT)-dependent photoperiodic pathway in lotus japonicus I: Verification of the flowering-associated function of an FT homolog. Biosci Biotechnol Biochem 77 (4):747-753.

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene VRN3 is an orthologue of FT. Proceedings of the National Academy of Sciences 103 (51):19581-19586.

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303 (5664):1640-1644.

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene VRN1. Proceedings of the National Academy of Sciences 100 (10):6263-6268.

Yang G, Zhai H, Wu HY, Zhang XZ, Lu SX, Wang YY, Li YQ, Hu B, Wang L, Wen ZX, Wang DC, Wang SD, Harada K, Xia ZJ, Xie FT (2017) QTL effects and epistatic interaction for flowering time and branch number in a soybean mapping population of JapanesexChinese cultivars. J Integr Agric 16 (9):1900-1912.

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. Plant Cell 12 (12):2473-2484.

Yoo SK, Wu X, Lee JS, Ahn JH (2011) AGAMOUS-LIKE 6 is a floral promoter that negatively regulates the FLC/MAF clade genes and positively regulates FT in Arabidopsis. Plant J 65 (1):62-76.

Young ND, Bharti AK (2012) Genome-enabled insights into legume biology. Annu Rev Plant Biol 63:283-305.

Yu Y, Liu Z, Wang L, Kim SG, Seo PJ, Qiao M, Wang N, Li S, Cao X, Park CM, Xiang F (2016) WRKY71 accelerates flowering via the direct activation of FLOWERING LOCUS T and LEAFY in Arabidopsis thaliana. Plant J 85 (1):96-106.

Yuan W, Luo X, Li Z, Yang W, Wang Y, Liu R, Du J, He Y (2016) A cis cold memory element and a trans epigenome reader mediate Polycomb silencing of FLC by vernalization in Arabidopsis. Nature Genetics 48:1527.

Zagotta MT, Hicks KA, Jacobs CI, Young JC, Hangarter RP, Meeks-Wagner DR (1996) The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. Plant J 10 (4):691-702.

Zahn LM, Kong H, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, dePamphilis CW, Ma H (2005) The Evolution of the SEPALLATA Subfamily of MADS-Box Genes: A Preangiosperm Origin With Multiple Duplications Throughout Angiosperm History. Genetics 169 (4):2209-2223.

Zaiter H, Barakat S (1995) Flower and pod abortion in chickpea as affected by sowing date and cultivar. Canadian journal of plant science 75 (2):321-327.

Zhang B, Wang L, Zeng L, Zhang C, Ma H (2015) Arabidopsis TOE proteins convey a photoperiodic signal to antagonize CONSTANS and regulate flowering time. Genes Dev 29 (9):975-987.

Zhang H, Pala M, Oweis T, Harris H (2000) Water use and water-use efficiency of chickpea and lentil in a Mediterranean environment. Australian Journal of Agricultural Research 51 (2):295-304.

Zhang J, Liu G, Guo C, He Y, Li Z, Ning G, Shi X, Bao M (2011) The FLOWERING LOCUS T orthologous gene of Platanus acerifolia is expressed as alternatively spliced forms with distinct spatial and temporal patterns. Plant Biol (Stuttg) 13 (5):809-820. doi: 810.1111/j.1438-8677.2010.00432.x. Epub 02011 Feb 00435.

Zhang X, Scheuring CF, Zhang M, Dong JJ, Zhang Y, Huang JJ, Lee MK, Abbo S, Sherman A, Shtienberg D, Chen W, Muehlbauer F, Zhang HB (2010) A BAC/BIBAC-based physical map of chickpea, Cicer arietinum L. BMC Genomics 11 (1).

Zhang X, Zhai H, Wang Y, Tian X, Zhang Y, Wu H, Lu S, Yang G, Li Y, Wang L, Hu B, Bu Q, Xia Z (2016) Functional conservation and diversification of the soybean maturity gene E1 and its homologs in legumes. Sci Rep 6:29548.

Zhang Y, Wang L (2005) The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evolutionary Biology 5 (1):1.

Zhao C, Takeshima R, Zhu J, Xu M, Sato M, Watanabe S, Kanazawa A, Liu B, Kong F, Yamada T, Abe J (2016) A recessive allele for delayed flowering at the soybean maturity locus E9 is a leaky allele of FT2a, a FLOWERING LOCUS T ortholog. Bmc Plant Biol 16:20.

Zhu M, Zhao S (2007) Candidate Gene Identification Approach: Progress and Challenges. International Journal of Biological Sciences 3 (7):420-427.

Zohary D, Hopf M (2000) Domestication of plants in the Old World. 3rd edn. 316pp. New York: Oxford University Press.

Appendix 3.1 Accession number of the 234 *Arabidopsis thaliana* flowering-related genes used for homolog-searching in *Cicer arietinum*. Further information about homologs found for each group can be found in the indicated appendix column.

Gene symbol	Gene name	Accession	Appendix
Phosphatidylethanolamine-binding protein (PEBP) family			
FT	*FLOWERING LOCUS T*	AT1G65480	3.21
TSF	*TWIN SISTER OF FT*	AT4G20370	
CEN	*CENTRORADIALIS*	AT2G27550	
TFL1	*TERMINAL FLOWER 1*	AT5G03840	
MFT	*MOTHER OF FT*	AT1G18100	
BFT	*BROTHER OF FT*	AT5G62040	
***CONSTANS* and *CONSTANS-like* genes**			
CO	*CONSTANS*	AT5G15840	3.19
COL1	*CONSTANS-LIKE 1*	AT5G15850	
COL2	*CONSTANS-LIKE 2*	AT3G02380	
COL3	*CONSTANS-LIKE 3*	AT2G24790	
COL4	*CONSTANS-LIKE 4*	AT5G24930	
COL5	*CONSTANS-LIKE 5*	AT5G57660	
COL6	*CONSTANS-LIKE 6*	AT1G68520	
COL7	*CONSTANS-LIKE 7*	AT1G73870	
COL8	*CONSTANS-LIKE 8*	AT1G49130	
COL9	*CONSTANS-LIKE 9*	AT3G07650	
COL10	*CONSTANS-LIKE 10*	AT5G48250	
COL11	*CONSTANS-LIKE 11*	AT4G15250	
COL12	*CONSTANS-LIKE 12*	AT3G21880	
COL13	*CONSTANS-LIKE 13*	AT2G47890	
COL14	*CONSTANS-LIKE 14*	AT2G33500	
COL15	*CONSTANS-LIKE 15*	AT1G28050	
COL16	*CONSTANS-LIKE 16*	AT1G25440	
Floral promoter and repressor			
FPF1	*FLOWERING PROMOTING FACTOR 1*	AT5G24860	3.26
FPF1-like	*FPF1-like*	AT4G31380	
FPF1-like	*FPF1-like*	AT5G10625	
TOE1	See other flower identity genes	ATEG28550	
TFL2	*TERMINAL FLOWER 2*	AT5G17690	3.27
EMF1	See other epigenetic regulators	AT5G11530	
EMF2	See other epigenetic regulators	AT5G51230	
HOS1	*HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1*	AT2G39810	3.28
ELF5	*EARLY FLOWERING 5*	AT5G62640	3.29
ELF6	*EARLY FLOWERING 6*	AT5G04240	3.30
REF6	*RELATIVE OF EARLY FLOWERING 6*	AT3G48430	3.31

Autonomous pathway

FCA	*FLOWERING TIME CONTROL PROTEIN FCA ALPHA*	AT4G16280	3.11
FY	*FY*	AT5G13480	
FLD	*FLOWERING LOCUS D*	AT3G10390	
FVE	*FVE*	AT2G19520	
FPA	*FPA*	AT2G43410	
FLK	*FLOWERING LOCUS K*	AT3G04610	
LD	*LUMINIDEPENDENS*	AT4G02560	
FWA	*FLOWERING WAGENINGEN*	AT4G25530	

Appendix 3.1 Continued

Gene symbol	Gene name	Accession	Appendix
MADS (MIKC)			
AP1	*APETALA 1*	AT1G69120	3.2
CAL	*CAULIFLOWER*	AT1G26310	
FUL	*FRUITFULL*	AT5G60910	
SEP1	*SEPALLATA 1*	AT5G15800	
SEP2	*SEPALLATA 2*	AT3G02310	
SEP3	*SEPALLATA 3*	AT1G24260	
SEP4	*SEPALLATA 4*	AT2G03710	
SOC1	*SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1*	AT2G45660	
SVP	*SHORT VEGETATIVE PHASE*	AT2G22540	
AP3	*APETALA 3*	AT3G54340	
PI	*PISTILLATA*	AT5G20240	
AG	*AGAMOUS*	AT4G18960	
SHP1	*SHATTERPROOF 1*	AT3G58780	
SHP2	*SHATTERPROOF 2*	AT2G42830	
STK	*SEEDSTICK*	AT4G09960	
ANR1	*ARABIDOPSIS NITRATE REGULATED 1*	AT2G14210	
FLC	*FLOWERING LOCUS C*	AT5G61920	
MAF1	*MADS AFFECTING FLOWERING-1*	AT1G77080	
MAF2	*MADS AFFECTING FLOWERING-2*	AT5G65050	
MAF3	*MADS AFFECTING FLOWERING-3*	AT5G65060	
MAF4	*MADS AFFECTING FLOWERING-4*	AT5G65070	
MAF5	*MADS AFFECTING FLOWERING-5*	AT5G65080	
TT16	*TRANSPARENT TESTA 16*	AT5G23260	
GOA	*GORDITA*	AT1G31140	
AGL6	*AGAMOUS-like 6*	AT2G45650	
AGL12	*AGAMOUS-like 12*	AT1G71692	
AGL13	*AGAMOUS-like 13*	AT3G61120	
AGL14	*AGAMOUS-like 14*	AT4G11880	
AGL15	*AGAMOUS-like 15*	AT5G13790	
AGL16	*AGAMOUS-like 16*	AT3G57230	
AGL17	*AGAMOUS-like 17*	AT2G22630	
AGL18	*AGAMOUS-like 18*	AT3G57390	
AGL19	*AGAMOUS-like 19*	AT4G22950	
AGL21	*AGAMOUS-like 21*	AT4G37940	
AGL24	*AGAMOUS-like 24*	AT4G24540	
AGL30	*AGAMOUS-like 30*	AT2G03060	
AGL33	*AGAMOUS-like 15*	AT2G26320	
AGL42	*AGAMOUS-like 42*	AT5G62165	
AGL65	*AGAMOUS-like 65*	AT1G18750	
AGL66	*AGAMOUS-like 66*	AT1G77980	
AGL67	*AGAMOUS-like 67*	AT1G77950	
AGL71	*AGAMOUS-like 71*	AT5G51870	
AGL72	*AGAMOUS-like 72*	AT5G51860	
AGL79	*AGAMOUS-like 79*	AT3G30260	
AGL94	*AGAMOUS-like 94*	AT1G69540	
AGL104	*AGAMOUS-like 104*	AT1G22130	

Appendix 3.1 Continued

Gene symbol	Gene name	Accession	Appendix
Phase change pathway			
SPL1	*SQUAMOSA PROMOTER-BINDING-like 1*	AT2G47070	3.12
SPL2	*SQUAMOSA PROMOTER-BINDING-like 2*	AT5G43270	
SPL3	*SQUAMOSA PROMOTER-BINDING-like 3*	AT2G33810	
SPL4	*SQUAMOSA PROMOTER-BINDING-like 4*	AT1G53160	
SPL5	*SQUAMOSA PROMOTER-BINDING-like 5*	AT3G15270	
SPL6	*SQUAMOSA PROMOTER-BINDING-like 6*	AT1G69170	
SPL7	*SQUAMOSA PROMOTER-BINDING-like 7*	AT5G18830	
SPL8	*SQUAMOSA PROMOTER-BINDING-like 8*	AT1G02065	
SPL9	*SQUAMOSA PROMOTER-BINDING-like 9*	AT2G42200	
SPL10	*SQUAMOSA PROMOTER-BINDING-like 10*	AT1G27370	
SPL11	*SQUAMOSA PROMOTER-BINDING-like 11*	AT1G27360	
SPL12	*SQUAMOSA PROMOTER-BINDING-like 12*	AT3G60030	
SPL13A	*SQUAMOSA PROMOTER-BINDING-like 13A*	AT5G50570	
SPL13B	*SQUAMOSA PROMOTER-BINDING-like 13B*	AT5G50670	
SPL14	*SQUAMOSA PROMOTER-BINDING-like 14*	AT1G20980	
SPL15	*SQUAMOSA PROMOTER-BINDING-like 15*	AT3G57920	
SQN	*SQUINT*	AT2G15790	3.13
HST	*HASTY*	AT3G05040	3.14
HYL1	*HYPONASTIC LEAVES 1*	AT1G09700	
DCL1	*DICER-like 1*	AT1G01040	3.15
DCL2	*DICER-like 2*	AT3G03300	
DCL3	*DICER-like 3*	AT3G43920	
DCL4	*DICER-like 4*	AT5G20320	
AGO1	*ARGONAUTE 1*	AT1G48410	3.16
AGO2	*ARGONAUTE 2*	AT1G31280	
AGO3	*ARGONAUTE 3*	AT1G31290	
AGO4	*ARGONAUTE 4*	AT2G27040	
AGO5	*ARGONAUTE 5*	AT2G27880	
AGO6	*ARGONAUTE 6*	AT2G32940	
AGO7	*ARGONAUTE 7*	AT1G69440	
AGO8	*ARGONAUTE 8*	AT5G21030	
AGO9	*ARGONAUTE 9*	AT5G21150	
AGO10	*ARGONAUTE 10*	AT5G43810	
Photoperiod pathway			
PFT1	*PHYTOCHROME AND FLOWERING TIME 1*	AT1G25540	3.4
FE (APL)	*FE*	AT1G79430	
FD	*FD*	AT4G35900	
FDP	*FD PARALOG*	AT2G17770	
VOZ1	*VASCULAR PLANT ONE ZINC FINGER 1*	AT1G28520	
VOZ2	*VASCULAR PLANT ONE ZINC FINGER 2*	AT2G42400	
TEM1	*TEMPRANILLO 1*	AT1G25560	
TEM2	*TEMPRANILLO 2*	AT1G68840	
FTIP1	*FT-INTERACTING PROTEIN 1*	AT5G06850	
DNF	*DAY NEUTRAL FLOWERING*	AT3G19140	

Appendix 3.1 Continued

Gene symbol	Gene name	Accession	Appendix
Vernalisation pathway			
VRN1	*VERNALIZATION 1*	AT3G18990	3.18
VIN3	*VERNALIZATION-INSENSITIVE 3*	AT5G57380	
FRI	*FRIGIDA*	AT4G00650	
FRL1	*FRIGIDA-like 1*	AT5G16320	
FRL2	*FRIGIDA-like 2*	AT1G31814	
VIP3	*VERNALIZATION-INDEPENDENCE 3*	AT4G29830	
VIP4	*VERNALIZATION-INDEPENDENCE 4*	AT5G61150	
ESD4	*EARLY IN SHORT DAYS 4*	AT4G15880	
PIE1	*PHOTOPERIOD-INDEPENDENT EARLY FLOWERING 1*	AT3G12810	
VRN2	See other epigenetic regulators	AT4g16845	
Clock genes			
ELF3	*EARLY FLOWERING 3*	AT2G25930	3.5
ELF4	*EARLY FLOWERING 4*	AT2G40080	
GI	*GIGANTEA*	AT1G22770	3.6
LHY	LATE ELONGATED HYPOCOTYL	AT1G01060	3.7
CCA1	*CIRCADIAN CLOCK ASSOCIATED 1*	AT2G46830	3.8
LCL1	*LHY/CCA1-like 1*	AT5G02840	
RVE8	*REVEILLE 8*	AT3G09600	
CHE	*TCP TF*	AT5G08330	
LUX	*LUX ARRHYTHMO*	AT3G46640	
Boa	*BROTHER OF LUX ARRHYTHMO*	AT5G59570	
LWD1	*LIGHT-REGULATED WD 1*	AT1G12910	
LWD2	*LIGHT-REGULATED WD 2*	AT3G26640	
LNK1	*NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED 1*	AT5G64170	
LNK2	*NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED 2*	AT3G54500	
TIC	*TIME FOR COFFEE*	AT3G22380	
PARG1/TEJ	*Poly(ADP-ribose) glycohydrolase*	AT2G31870	
CDF1	*CYCLING DOF FACTOR 1*	AT5G62430	3.9
CDF2	*CYCLING DOF FACTOR 2*	AT5G39660	
CDF3	*CYCLING DOF FACTOR 3*	AT3G47500	
CDF4	*CYCLING DOF FACTOR 4*	AT2G34140	
CDF5	*CYCLING DOF FACTOR 5*	AT1G69570	
Other flower identity genes			
LFY	*LEAFY*	AT5G61850	3.32
UFO	*UNUSUAL FLORAL ORGANS*	AT1G30950	3.33
AP2	*APETALA 2*	AT4G36920	3.34
TOE1	*TARGET OF EARLY ACTIVATION TAGGED 1*	AT2G28550	
TOE2	*TARGET OF EARLY ACTIVATION TAGGED 2*	AT5G60120	
TOE3	*TARGET OF EARLY ACTIVATION TAGGED 3*	AT5G67180	
SNZ	*SCHNARCHZAPFEN*	AT2G39250	
SMZ	*SCHLAFMUTZE*	AT3G54990	
Legume specific			
E1	*E1*	Medtr1g045740	3.4

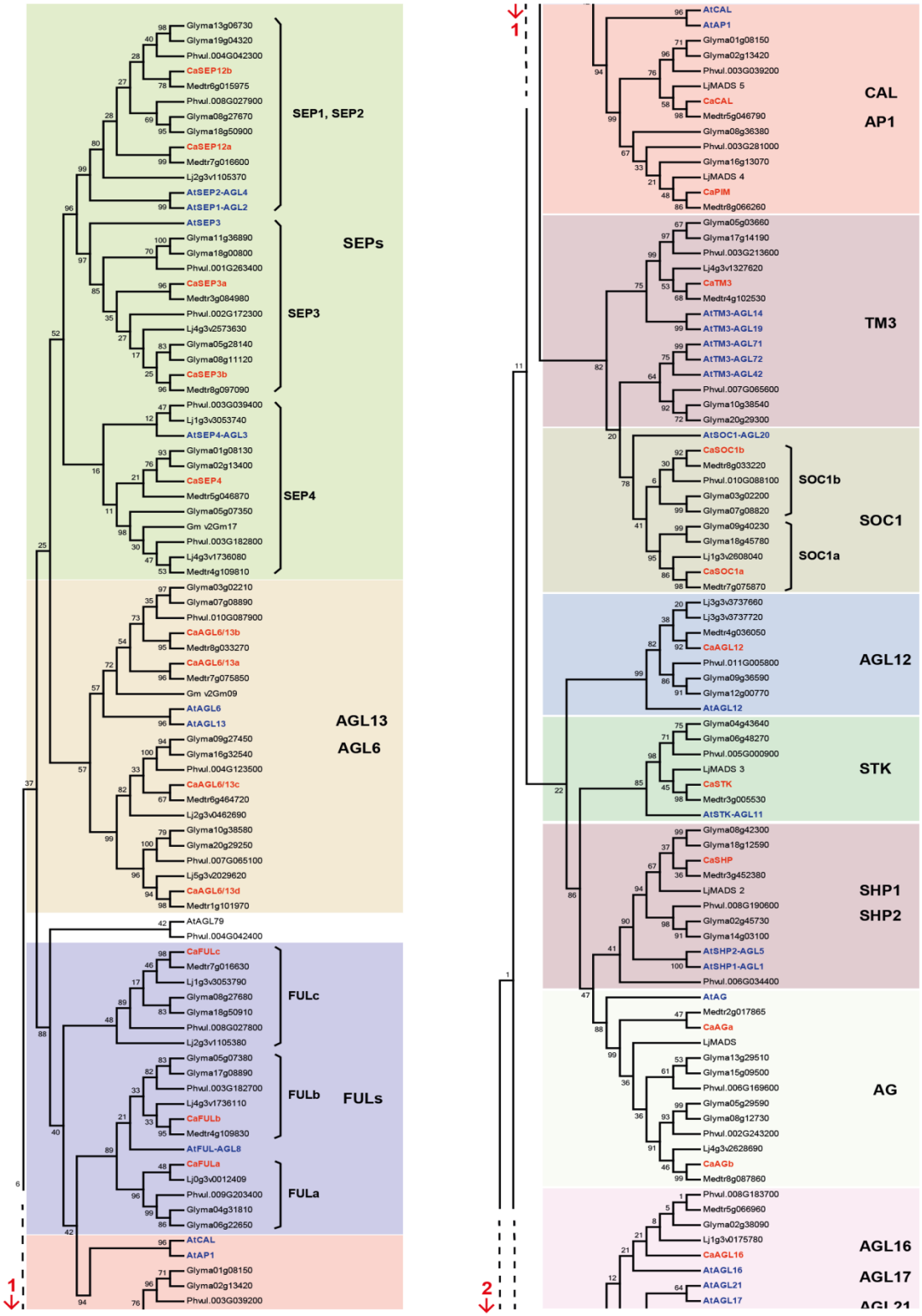
Appendix 3.1 Continued

Gene symbol	Gene name	Accession	Appendix
Gibberellin oxidases			
GA2ox1	*GIBBERELLIN 2-OXIDASE 1*	AT1G78440	3.17
GA2ox2	*GIBBERELLIN 2-OXIDASE 2*	AT1G30040	
GA2ox3	*GIBBERELLIN 2-OXIDASE 3*	AT2G34550	
GA2ox4	*GIBBERELLIN 2-OXIDASE 4*	AT1G47990	
GA2ox6	*GIBBERELLIN 2-OXIDASE 6*	AT1G02400	
GA2ox7	*GIBBERELLIN 2-OXIDASE 7*	AT1G50960	
GA2ox8	*GIBBERELLIN 2-OXIDASE 8*	AT4G21200	
GA3ox1	*GIBBERELLIN 3-OXIDASE 1*	AT1G15550	
GA3ox2 (GA4H)	*GIBBERELLIN 3-OXIDASE 2*	AT1G80340	
GA3ox3	*GIBBERELLIN 3-OXIDASE 3*	AT4G21690	
GA3ox4	*GIBBERELLIN 3-OXIDASE 4*	AT1G80330	
GA20ox1	*GIBBERELLIN 20-OXIDASE 1*	AT4G25420	
GA20ox2	*GIBBERELLIN 20-OXIDASE 2*	AT5G51810	
GA20ox3	*GIBBERELLIN 20-OXIDASE 3*	AT5G07200	
GA20ox4	*GIBBERELLIN 20-OXIDASE 4*	AT1G60980	
GA20ox5	*GIBBERELLIN 20-OXIDASE 5*	AT1G44090	
Pseudo-response regulators			
TOC1	*TIMING OF CAB EXPRESSION 1*	AT5G61380	3.10
PRR3	*PSEUDO RESPONSE REGULATOR 3*	AT5G60100	
PRR5	*PSEUDO RESPONSE REGULATOR 5*	AT5G24470	
PRR7	*PSEUDO RESPONSE REGULATOR 7*	AT5G02810	
PRR9	*PSEUDO RESPONSE REGULATOR 9*	AT2G46790	
Other epigenetic regulators			
ATX1	*HOMOLOGUE OF TRITHORAX 1*	AT2G31650	3.35
ATX2	*HOMOLOGUE OF TRITHORAX 2*	AT1G05830	
ATX3	*HOMOLOGUE OF TRITHORAX 3*	AT3G61740	
ATX4	*HOMOLOGUE OF TRITHORAX 4*	AT4G27910	
ATX5	*HOMOLOGUE OF TRITHORAX 5*	AT5G53430	
EFS/SDG8	*EARLY FLOWERING IN SHORT DAYS*	AT1G77300	3.36
ELF7	*EARLY FLOWERING 7*	AT1G79730	3.37
HUB1	*HISTONE MONOUBIQUITINATION 1*	AT2G44950	3.38
HUB2	*HISTONE MONOUBIQUITINATION 2*	AT1G55250	
CLF	*CURLY LEAF*	AT2G23380	3.39
EMF1	*EMBRYONIC FLOWER 1*	AT5G11530	
EMF2	*EMBRYONIC FLOWER 2*	AT5G51230	
FIE(FIS3)	*FERTILIZATION-INDEPENDENT ENDOSPERM*	AT3G20740	
FIS2	*FERTILIZATION INDEPENDENT SEED 2*	AT2G35670	
MEA (MEDEA, FIS1)	*MALE-ENHANCED ANTIGEN*	AT1G02580	
MSI1	*MALE STERILITY1*	AT5G58230	
SWN (EZA1)	*SWINGER*	AT4G02020	
VRN2	*VERNALIZATION 2*	AT4G16845	

Appendix 3.1 Continued

Gene symbol	Gene name	Accession	Appendix
Photoreceptors and light signalling			
PHYA	*PHYTOCHROME A*	AT1G09570	
PHYB	*PHYTOCHROME B*	AT2G18790	
PHYC	*PHYTOCHROME C*	AT5G35840	
PHYD	*PHYTOCHROME D*	AT4G16250	
PHYE	*PHYTOCHROME E*	AT4G18130	
CRY1	*CRYPTOCHROME 1*	AT4G08920	
CRY2	*CRYPTOCHROME 2*	AT1G04400	
FKF1	*FLAVIN-BINDING, KELCH REPEAT, F BOX 1*	AT1G68050	
ZTL	*ZEITLUPE*	AT5G57360	
LKP2	*LOV KELCH REPEAT PROTEIN 2*	AT2G18915	
PIF1	*PHYTOCHROME INTERACTING FACTOR 1*	AT2G20180	
PIF3	*PHYTOCHROME INTERACTING FACTOR 3*	AT1G09530	3.3
PIF4	*PHYTOCHROME INTERACTING FACTOR 4*	AT2G43010	
PIF5	*PHYTOCHROME INTERACTING FACTOR 5*	AT3G59060	
PIF6	*PHYTOCHROME INTERACTING FACTOR 6*	AT3G62090	
PIF7	*PHYTOCHROME INTERACTING FACTOR 7*	AT5G61270	
PIF8	*PHYTOCHROME INTERACTING FACTOR 8*	AT4G00050	
PIL1	*PIL1*	AT2G46970	
ALC	*ALC*	AT5G67110	
SPT	*SPT*	AT4G36930	
COP1	*CONSTITUTIVE PHOTOMORPHOGENIC 1*	AT2G32950	
SPA1	*SUPRESSOR OF PHYA 1*	AT2G46340	
SPA3	*SUPRESSOR OF PHYA 3*	AT3G15354	
SPA4	*SUPRESSOR OF PHYA 4*	AT1G53090	

Appendix 3.2 Maximum likelihood tree derived of the alignment of the type II (MIKC) MADS genes from *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Medtr), *Phaseolus vulgaris* (Phvul), *Glycine max* (Gm, Glyma), *Lotus japonicus* (Lj) and *Arabidopsis thaliana* (At, in blue). Protein sequences from accessions listed in Table 1 were aligned using MUSCLE and tree construction were performed using MEGA6 software. Numbers in branches represent bootstrap values from 1000 replications.



Appendix 3.2 Continued

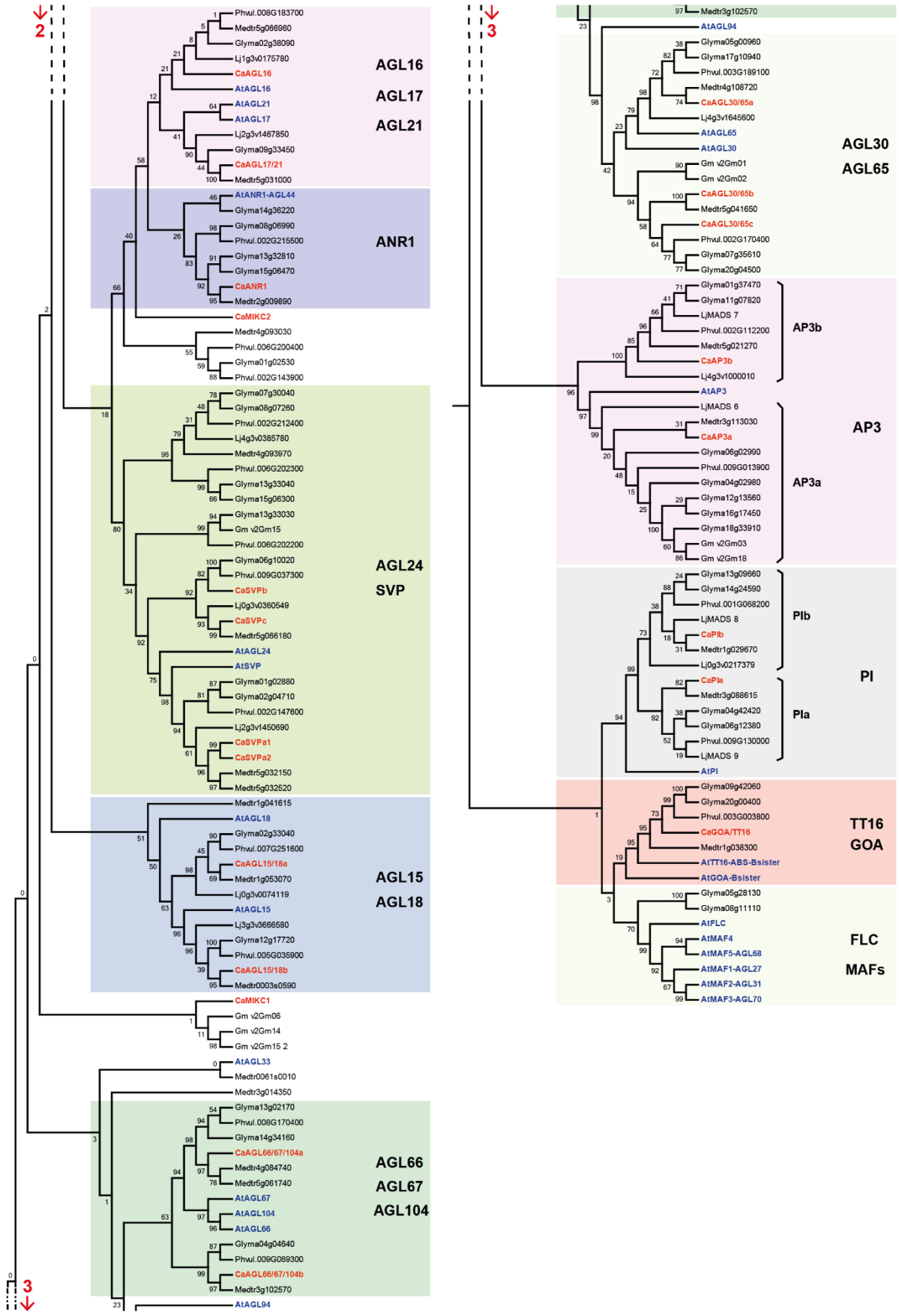


Table 1 Accession numbers for the MADS-MIKC protein sequences used in the alignment and tree. Accessions from *Arabidopsis thaliana* and *Glycine max* were obtained from (Gramzow and Theißen 2013) and those from *Lotus japonicus* lacking number from Dong et al. (2005).

Phaseolus vulgaris
Phvul.001G068200
Phvul.001G263400
Phvul.002G112200
Phvul.002G143900
Phvul.002G147600
Phvul.002G170400
Phvul.002G172300
Phvul.002G212400
Phvul.002G215500
Phvul.002G243200
Phvul.003G003800
Phvul.003G039200
Phvul.003G039400
Phvul.003G182700
Phvul.003G182800
Phvul.003G189100
Phvul.003G213600
Phvul.003G281000
Phvul.004G042300
Phvul.004G042400
Phvul.004G123500
Phvul.005G000900
Phvul.005G035900
Phvul.006G034400
Phvul.006G169600
Phvul.006G200400
Phvul.006G202200
Phvul.006G202300
Phvul.007G065100
Phvul.007G065600
Phvul.007G251600
Phvul.008G027800
Phvul.008G027900
Phvul.008G170400
Phvul.008G183700
Phvul.008G190600
Phvul.009G013900
Phvul.009G037300
Phvul.009G069300
Phvul.009G130000
Phvul.009G203400
Phvul.010G087900
Phvul.010G088100
Phvul.011G005800

Cicer arietinum	
CaPIM	LOC101488241
CaCAL	LOC101503680
CaFULa	LOC101510419
CaFULb	LOC101507855
FULc	LOC101513562
CaSEP12a	LOC101513013
CaSEP12b	LOC101508487
CaSEP3a	LOC101503539
CaSEP3b	LOC101501925
CaSEP4	LOC101504006
CaAGL6/13a	LOC101510444
CaAGL6/13b	LOC101490010
CaAGL6/13c	LOC101491642
CaAGL6/13d	LOC101488623
CaSOC1a	LOC101510775
CaSOC1b	LOC101489691
CaSVPa1	LOC101497665
CaSVPa2	LOC101513905
CaSVPb	LOC101509413
CaSVPc	LOC101503022
CaAP3a	LOC101514275
CaAP3b	LOC101490232
CaPIa	LOC101502768
CaPIb	LOC101491763
CaAGa	LOC101493118
CaAGb	LOC101512426
CaSHP	LOC101493068
CaSTK	LOC101498947
CaANR1	LOC101500250
CaAGL16	LOC101509359
CaAGL17/21	LOC101509049
CaAGL12	LOC101508958
CaAGL15/18a	LOC101513885
CaAGL15/18b	LOC101502231
CaGOA/TT16	LOC101500959
CaAGL66/67/104a	LOC101497019
CaAGL66/67/104b	LOC101506754
CaAGL30/65a	LOC101502697
CaAGL30/65b	LOC101504656
CaAGL30/65c	LOC101510075
CaMIKC1	LOC101515233
CaMIKC2	LOC101491267

Lotus japonicus
Lj0g3v0012409
Lj0g3v0074119
Lj0g3v0217379
Lj0g3v0360549
Lj1g3v0175780
Lj1g3v2608040
Lj1g3v3053740
Lj1g3v3053790
Lj2g3v0462690
Lj2g3v1105370
Lj2g3v1105380
Lj2g3v1450690
Lj2g3v1467850
Lj3g3v3666580
Lj3g3v3737660
Lj3g3v3737720
Lj4g3v0385780
Lj4g3v1000010
Lj4g3v1327620
Lj4g3v1645600
Lj4g3v1736080
Lj4g3v1736110
Lj4g3v2573630
Lj4g3v2628690
Lj5g3v2029620
LjMADS
LjMADS 2
LjMADS 3
LjMADS 4
LjMADS 5
LjMADS 6
LjMADS 7
LjMADS 8
LjMADS 9

Table 1 Continued

Glycine max		*Arabidopsis thaliana*		*Medicago truncatula*
Glyma01g02530	Glyma10g38540	**AP1**	AT1G69120	Medtr1g029670
Glyma01g02880	Glyma10g38580	**CAL**	AT1G26310	Medtr1g038300
Glyma01g08130	Glyma11g07820	**FUL**	AT5G60910	Medtr1g041615
Glyma01g08150	Glyma11g36890	**SEP1**	AT5G15800	Medtr1g053070
Glyma01g37470	Glyma12g00770	**SEP2**	AT3G02310	Medtr1g101970
Glyma02g04710	Glyma12g13560	**SEP3**	AT1G24260	Medtr2g009890
Glyma02g13400	Glyma12g17720	**SEP4**	AT2G03710	Medtr2g017865
Glyma02g13420	Glyma13g02170	**AGL6**	AT2G45650	Medtr3g005530
Glyma02g33040	Glyma13g06730	**AGL13**	AT3G61120	Medtr3g014350
Glyma02g38090	Glyma13g09660	**SOC1**	AT2G45660	Medtr3g084980
Glyma02g45730	Glyma13g29510	**AGL14**	AT4G11880	Medtr3g088615
Glyma03g02200	Glyma13g32810	**AGL19**	AT4G22950	Medtr3g102570
Glyma03g02210	Glyma13g33030	**AGL42**	AT5G62165	Medtr3g113030
Glyma04g02980	Glyma13g33040	**AGL71**	AT5G51870	Medtr3g452380
Glyma04g04640	Glyma14g03100	**AGL72**	AT5G51860	Medtr0003s0590
Glyma04g31810	Glyma14g24590	**SVP**	AT2G22540	Medtr4g036050
Glyma04g42420	Glyma14g34160	**AGL24**	AT4G24540	Medtr4g084740
Glyma04g43640	Glyma14g36220	**AP3**	AT3G54340	Medtr4g093030
Glyma05g00960	Glyma15g06300	**PI**	AT5G20240	Medtr4g093970
Glyma05g03660	Glyma15g06470	**AG**	AT4G18960	Medtr4g102530
Glyma05g07350	Glyma15g09500	**SHP1**	AT3G58780	Medtr4g108720
Glyma05g07380	Glyma16g13070	**SHP2**	AT2G42830	Medtr4g109810
Glyma05g28130	Glyma16g17450	**STK**	AT4G09960	Medtr4g109830
Glyma05g28140	Glyma16g32540	**ANR1**	AT2G14210	Medtr5g021270
Glyma05g29590	Glyma17g08890	**AGL16**	AT3G57230	Medtr5g031000
Glyma06g02990	Glyma17g10940	**AGL17**	AT2G22630	Medtr5g032150
Glyma06g10020	Glyma17g14190	**AGL21**	AT4G37940	Medtr5g032520
Glyma06g12380	Glyma18g00800	**FLC**	AT5G61920	Medtr5g041650
Glyma06g22650	Glyma18g12590	**MAF1**	AT1G77080	Medtr5g046790
Glyma06g48270	Glyma18g33910	**MAF2**	AT5G65050	Medtr5g046870
Glyma07g08820	Glyma18g45780	**MAF3**	AT5G65060	Medtr5g061740
Glyma07g08890	Glyma18g50900	**MAF4**	AT5G65070	Medtr5g066180
Glyma07g30040	Glyma18g50910	**MAF5**	AT5G65080	Medtr5g066960
Glyma07g35610	Glyma19g04320	**AGL12**	AT1G71692	Medtr6g015975
Glyma08g06990	Glyma20g00400	**AGL15**	AT5G13790	Medtr6g464720
Glyma08g07260	Glyma20g04500	**AGL18**	AT3G57390	Medtr7g016600
Glyma08g11110	Glyma20g29250	**TT16**	AT5G23260	Medtr7g016630
Glyma08g11120	Glyma20g29300	**GOA**	AT1G31140	Medtr7g075850
Glyma08g12730	Gm_v2Gm01	**AGL104**	AT1G22130	Medtr7g075870
Glyma08g27670	Gm_v2Gm02	**AGL94**	AT1G69540	Medtr8g033220
Glyma08g27680	Gm_v2Gm03	**AGL67**	AT1G77950	Medtr8g033270
Glyma08g36380	Gm_v2Gm06	**AGL66**	AT1G77980	Medtr8g066260
Glyma08g42300	Gm_v2Gm09	**AGL65**	AT1G18750	Medtr8g087860
Glyma09g27450	Gm_v2Gm14	**AGL33**	AT2G26320	Medtr8g097090
Glyma09g33450	Gm_v2Gm15	**AGL30**	AT2G03060	Medtr0061s0010
Glyma09g36590	Gm_v2Gm15 2			
Glyma09g40230	Gm_v2Gm17			
Glyma09g42060	Gm_v2Gm18			

Appendix 3.3 Maximum parsimony tree derived from the alignment of genes involved in photoreception and light signalling in *Cicer arietinum* (Ca, red color), *Medicago truncatula* (Mt), *Pisum sativum* (Ps) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using MUSCLE and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications.

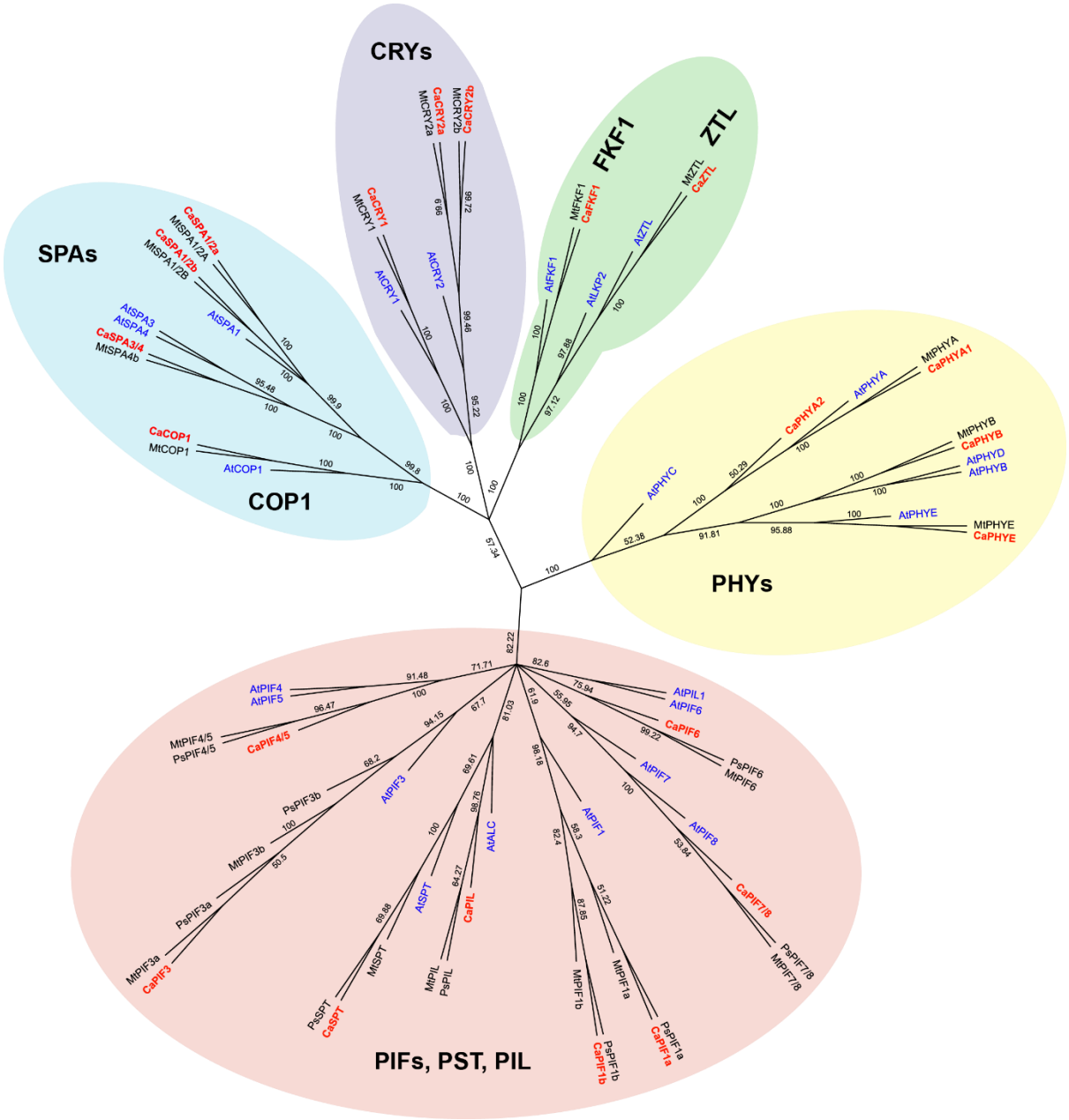


Table 1 Accession number of the sequences used in the alignment and tree construction of genes involved in photoreception and light signalling in four plant species.

Arabidopsis thaliana		***Medicago truncatula***		***Cicer arietinum***		***Pisum sativum***	
PHYA	AT1G09570	**PHYA**	Medtr1g085160	**PHYA1**	LOC101506511	**PHYA**	M37217
PHYB	AT2G18790	**PHYB**	Medtr2g034040	**PHYA2**	LOC101496082	**PHYB**	AF069305
PHYC	AT5G35840	**PHYE**	Medtr2g049520	**PHYB**	LOC101498651	**CRY1**	AY508969
PHYD	AT4G16250	**CRY1**	Medtr5g063920.1	**PHYE**	LOC101494841	**CRY2a**	AY508972
PHYE	AT4G18130	**CRY2a**	Medtr1g076190	**CRY1**	LOC101503257	**CRY2b**	AY508974
CRY1	AT4G08920	**CRY2b**	Medtr1g043180	**CRY2a**	LOC101509277	**PsPIF1a**	PsCam049573
CRY2	AT1G04400	**FKF1**	Medtr8g105590	**CRY2b**	LOC101497404	**PsPIF1b**	PsCam045632
FKF1	AT1G68050	**ZTL**	Medtr2g036510	**ZTL**	LOC101494213	**PsPIF3a**	PsCam045246
ZTL	AT5G57360	**PIF1a**	Medtr7g099540	**PIF1a**	LOC101512389	**PsPIF3b**	PsCam006747
LKP2	AT2G18915	**PIF1b**	Medtr1g069155	**PIF1b**	LOC101505427	**PsPIF4/5**	PsCam033784
PIF1	AT2G20180	**PIF3a**	Medtr1g084980	**PsPIF3**	LOC101501589	**PsPIF6**	PsCam042543
PIF3	AT1G09530	**PIF3b**	Medtr7g111320	**PIF4/5**	LOC101488979	**PsPIF7/8**	PsCam010797
PIF4	AT2G43010	**PIF4/5**	Medtr3g449770	**PIF6**	LOC101490374	**PsPIL**	GCML01008161
PIF5	AT3G59060	**PIF6**	Medtr7g110810	**PIF7/8**	LOC101492173	**PsSPT**	PsCam050219
PIF6	AT3G62090	**PIF7/8**	Medtr7g039110	**PIL**	LOC101503520		
PIL1	AT2G46970	**MtPIL**	Medtr1g019240	**PsSPT**	LOC101515727		
PIF7	AT5G61270	**MtSPT**	Medtr5g017040	**COP1**	LOC101510767		
PIF8	AT4G00050	**COP1**	Medtr5g085250	**SPA1/2a**	LOC101508100		
ALC	AT5G67110	**SPA1/2a**	Medtr5g009530	**SPA1/2b**	LOC101501801		
SPT	AT4G36930	**SPA1/2b**	Medtr8g027985	**SPA3/4**	LOC101488676		
COP1	AT2G32950	**SPA4a**	Medtr2g084980				
SPA1	AT2G46340	**SPA4b**	Medtr2g085210				
SPA3	AT3G15354	**SPA4c**	Medtr8g091170				
SPA4	AT1G53090						

Appendix 3.4 Maximum parsimony tree from the alignment of genes involved in the flowering photoperiod pathway in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt) and *Arabidopsis thaliana* (At, blue). Number in branches represent bootstrap consensus support from 1000 repetitions. Protein sequences from accessions in Table 1 were aligned using ClustalW and PAUP* software was used to build the tree. Numbers in branches represent bootstrap consensus support from 1000 replications.

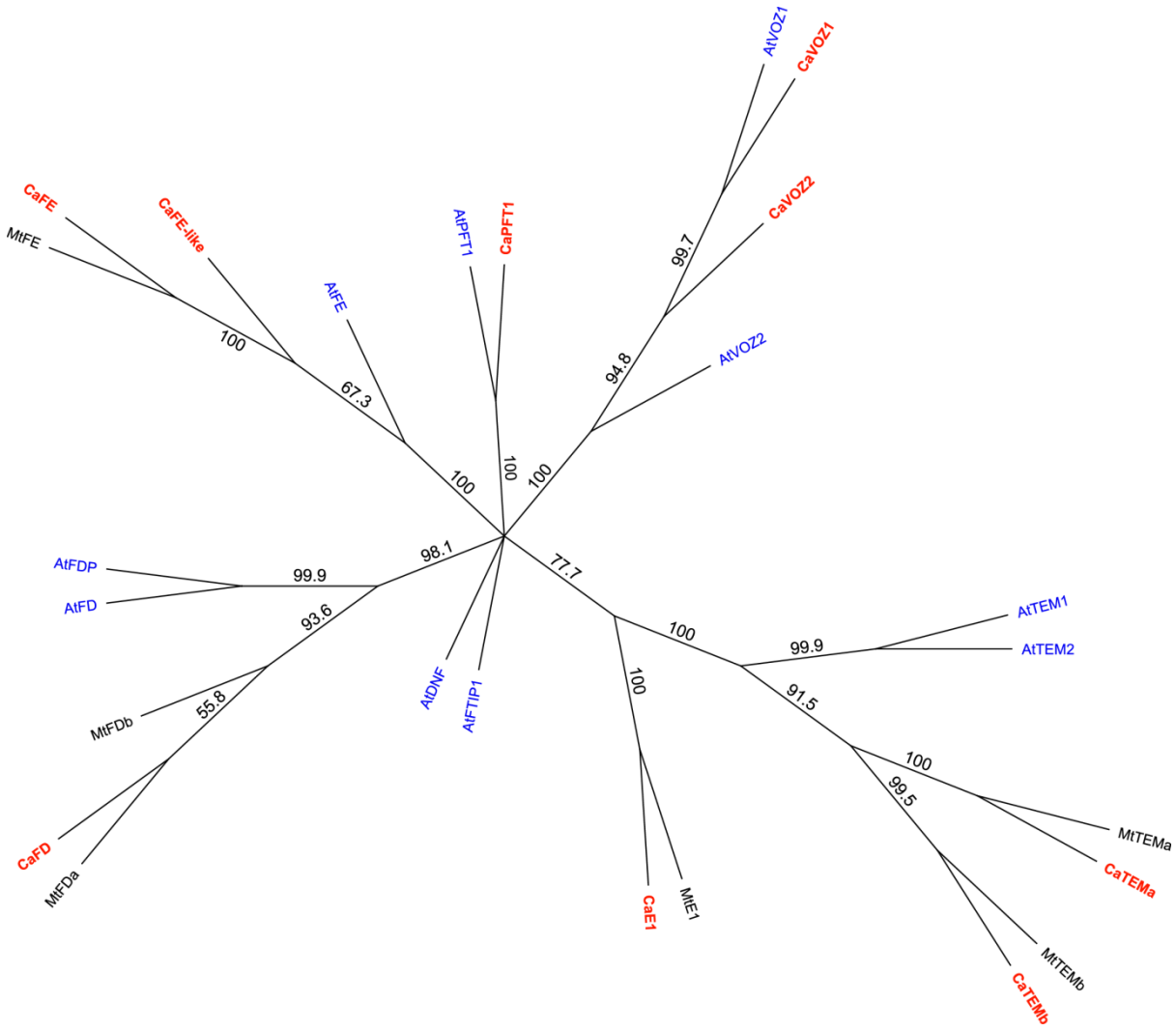


Table 1 Accession number of the genes used for the alignment and tree construction.

Arabidopsis thaliana	
PFT1	AT1G25540
FE	AT1G79430
FD	AT4G35900
FDP	AT2G17770
VOZ1	AT1G28520
VOZ2	AT2G42400
TEM1	AT1G25560
TEM2	AT1G68840
FTIP1	AT5G06850
DNF	AT3G19140

Medicago truncatula	
MtFE	Medtr6g444980
MtFDa	Medtr5g022780
MtFDb	Medtr8g075130
MtTEMa	Medtr5g053920.1
MtTEMb	Medtr1g093600
MtE1	Medtr2g058520

Cicer arietinum	
CaPFT1	LOC101507045
CaFE	LOC101507545
CaFE-like	LOC101504351
CaFD	LOC101513662
CaVOZ1	LOC101496785
CaVOZ2	LOC101494508
CaTEMa	LOC101503485
CaTEMb	LOC101492303
CaE1	LOC101497661

Appendix 3.5 Maximum parsimony tree derived from the alignment of *EARLY FLOWERING 3* (ELF3) and *EARLY FLOWERING 4* (ELF4) genes in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Pisum sativum* (Ps) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using ClustalW and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications.

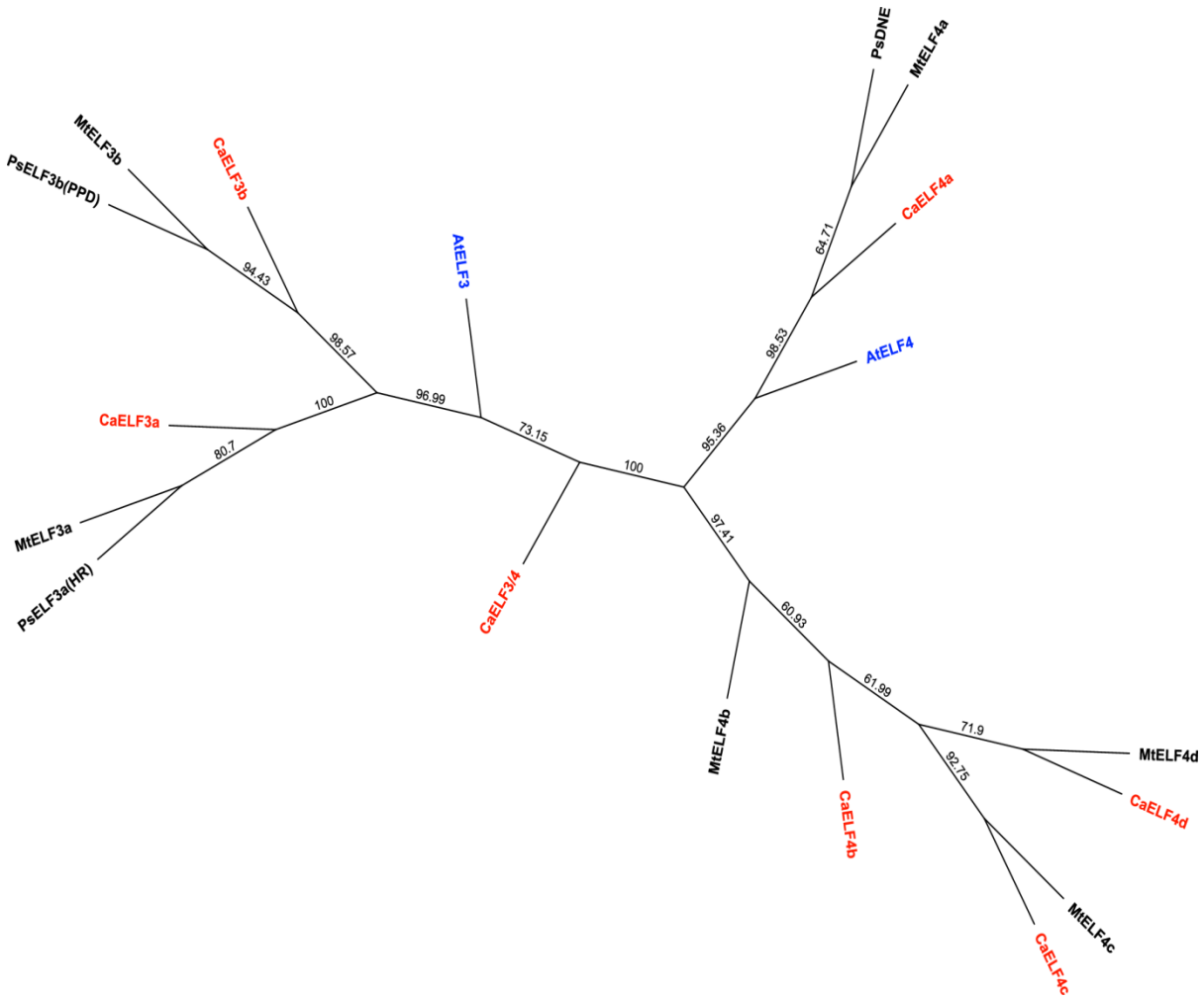


Table 1 Accession numbers of the sequences used in the alignment and tree construction of ELF3 and ELF4 genes in four plant species.

Arabidopsis thaliana	
ELF3	AT2G25930
ELF4	AT2G40080

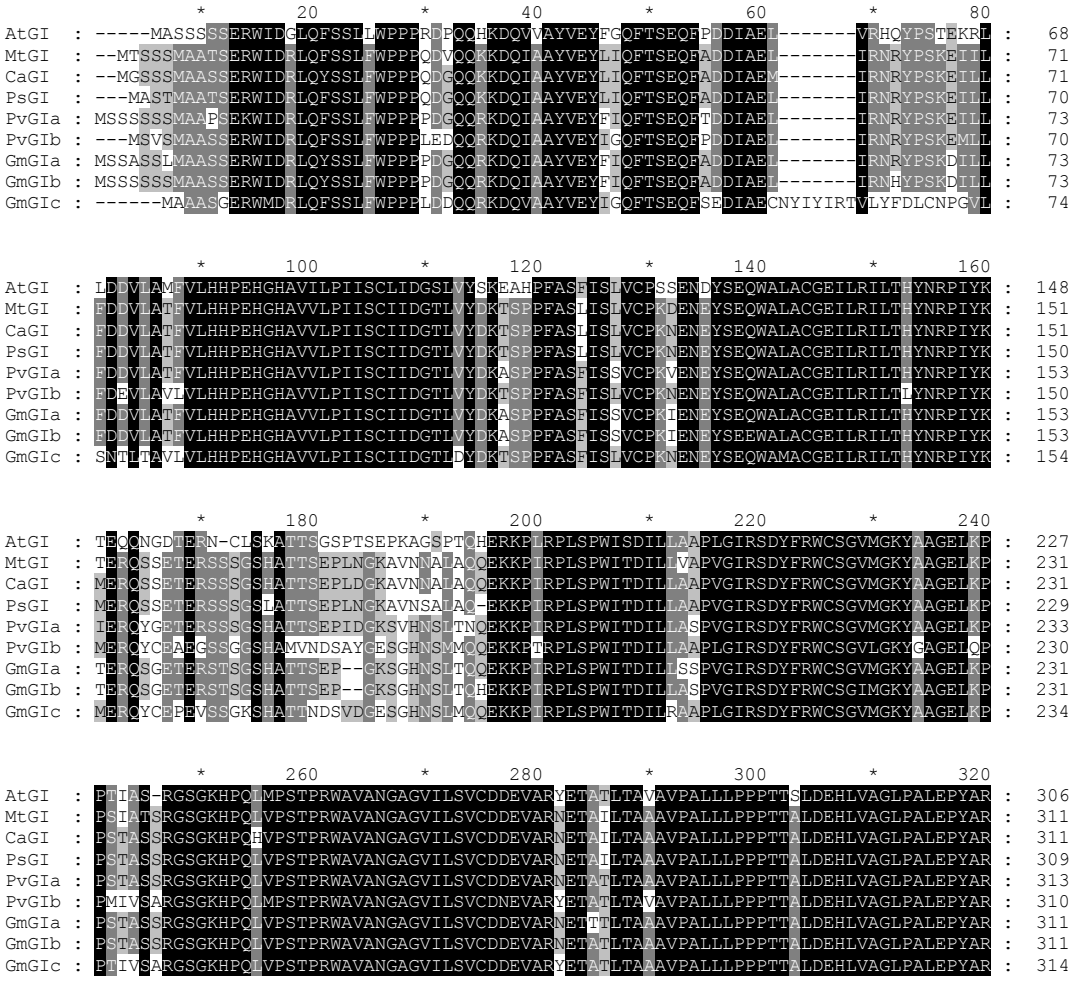
Cicer arietinum	
ELF3a	LOC101489432
ELF3b	LOC101488316
ELF34	LOC101495393
ELF4a	LOC101498412
ELF4b	LOC101506035
ELF4c	LOC101491318
ELF4d	LOC101504004

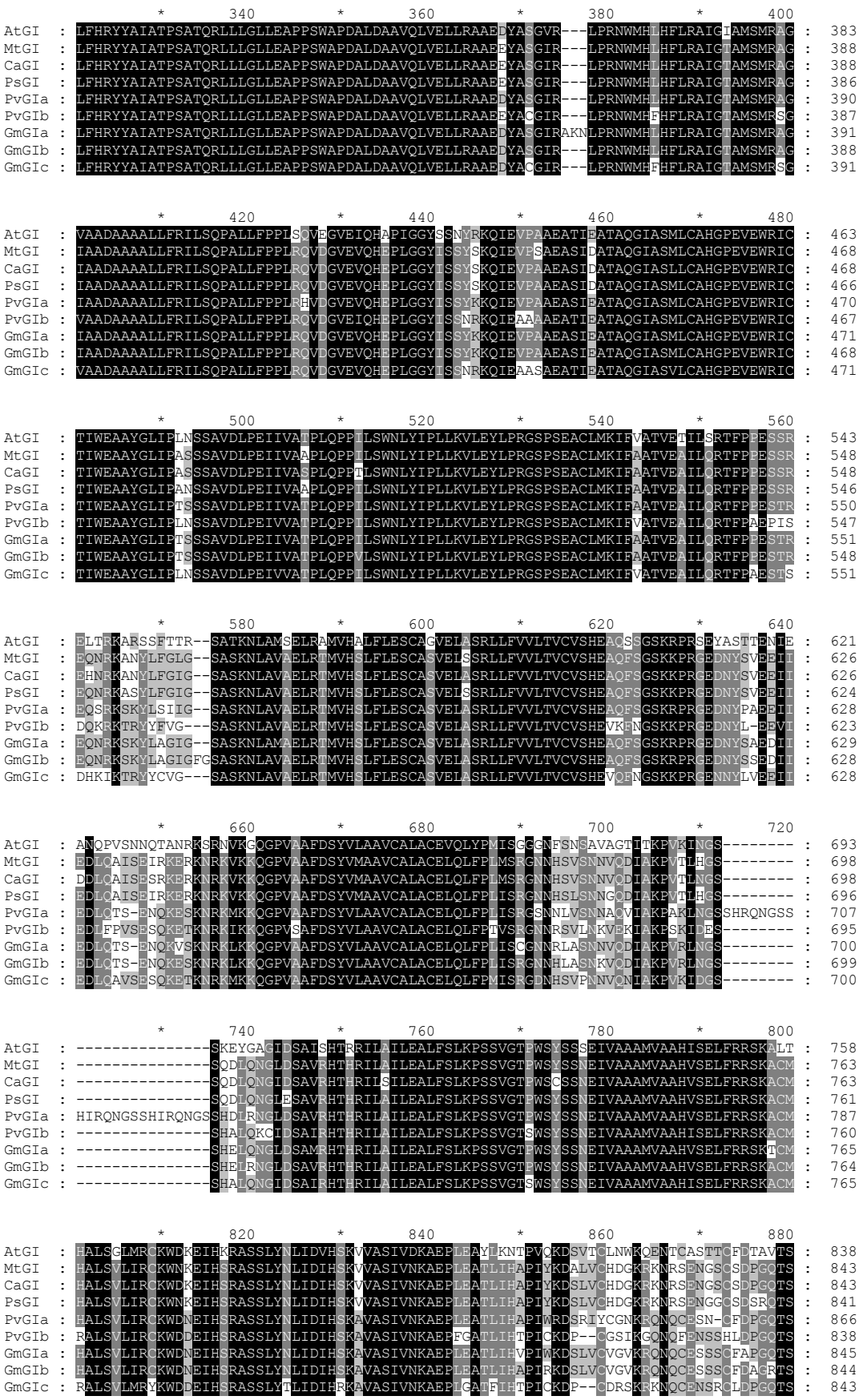
Medicago truncatula	
ELF3a	Medtr3g103970
ELF3b	Medtr1g016920
ELF4a	Medtr3g070490
ELF4b	Medtr4g125590
ELF4c	Medtr2g041310
ELF4d	Medtr8g020200

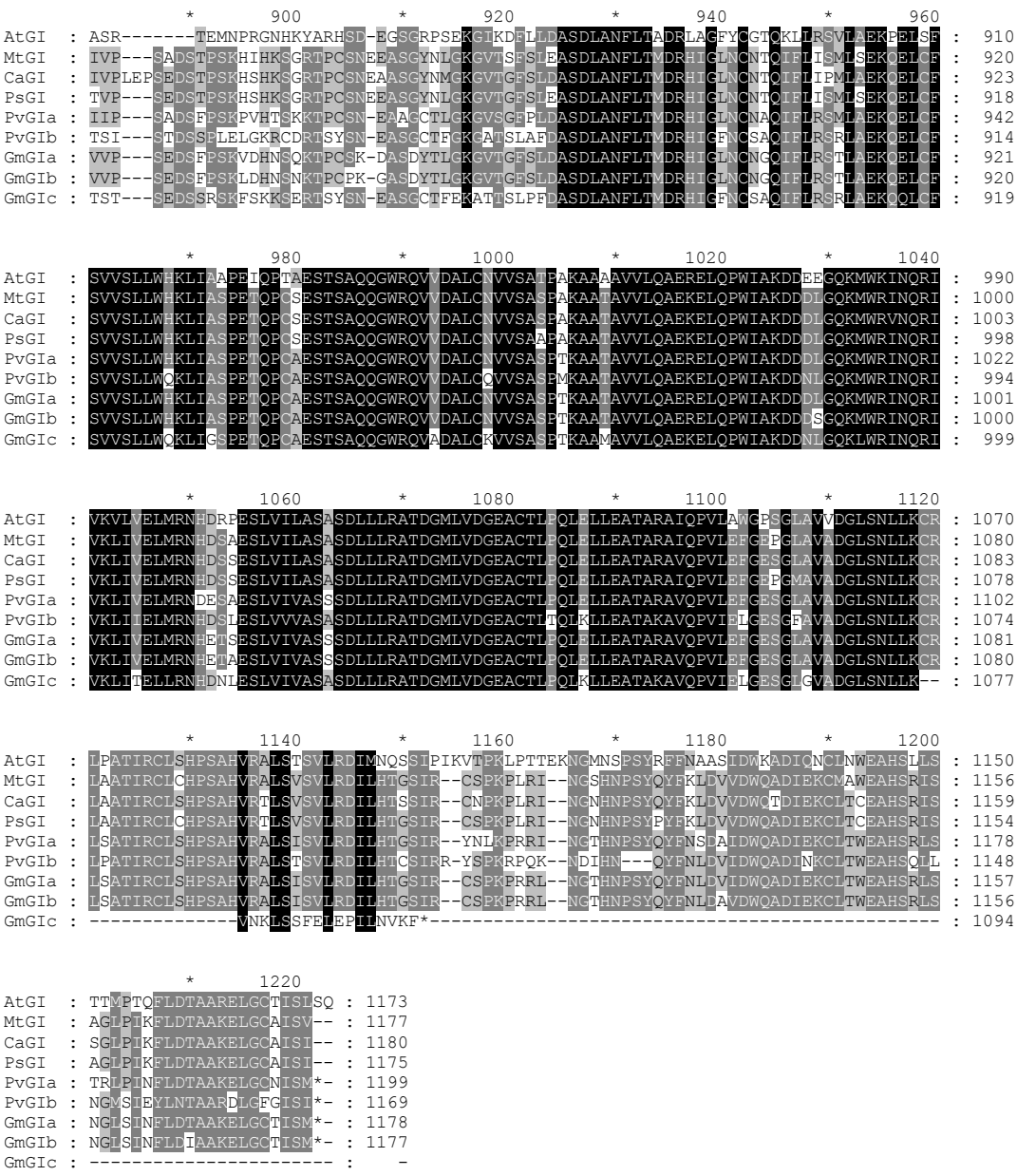
Pisum sativum	
HR	AFR60580
PPD	PsCam054737
DNE	AY830926

Appendix 3.6 Identity matrix (%) derived from the multiple sequence alignment of GIGANTEA proteins in the species *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt), *Cicer arietinum* (Ca), *Pisum sativum* (Ps), *Phaseolus vulgaris* (Pv) and *Glycine max* (Gm). Protein sequences from accessions listed in the table were aligned using ClustalW in Geneious 8 software.

	Accession	AtGI	MtGI	CaGI	PsGI	PvGIa	PvGIb	GmGIa	GmGIb	GmGIc
AtGI	AT1G22770		74.5	74.3	74.6	74.0	72.8	75.4	75.3	71.6
MtGI	Medtr1g098160	74.5		95.4	96.7	87.1	81.3	89.0	89.0	79.6
CaGI	LOC101511540	74.3	95.4		95.6	86.9	81.2	89.1	89.1	79.8
PsGI	EF185297	74.6	96.7	95.6		86.7	80.9	88.8	88.8	79.2
PvGIa	Phvul.007G083500	74.0	87.1	86.9	86.7		80.4	91.4	91.5	78.6
PvGIb	Phvul.004G088300	72.8	81.3	81.2	80.9	80.4		81.9	81.6	86.4
GmGIa	Glyma.20G170000	75.4	89.0	89.1	88.8	91.4	81.9		96.9	79.5
GmGIb	Glyma.10G221500	75.3	89.0	89.1	88.8	91.5	81.6	96.9		79.4
GmGIc	Glyma.16G163200	71.6	79.6	79.8	79.2	78.6	86.4	79.5	79.4	

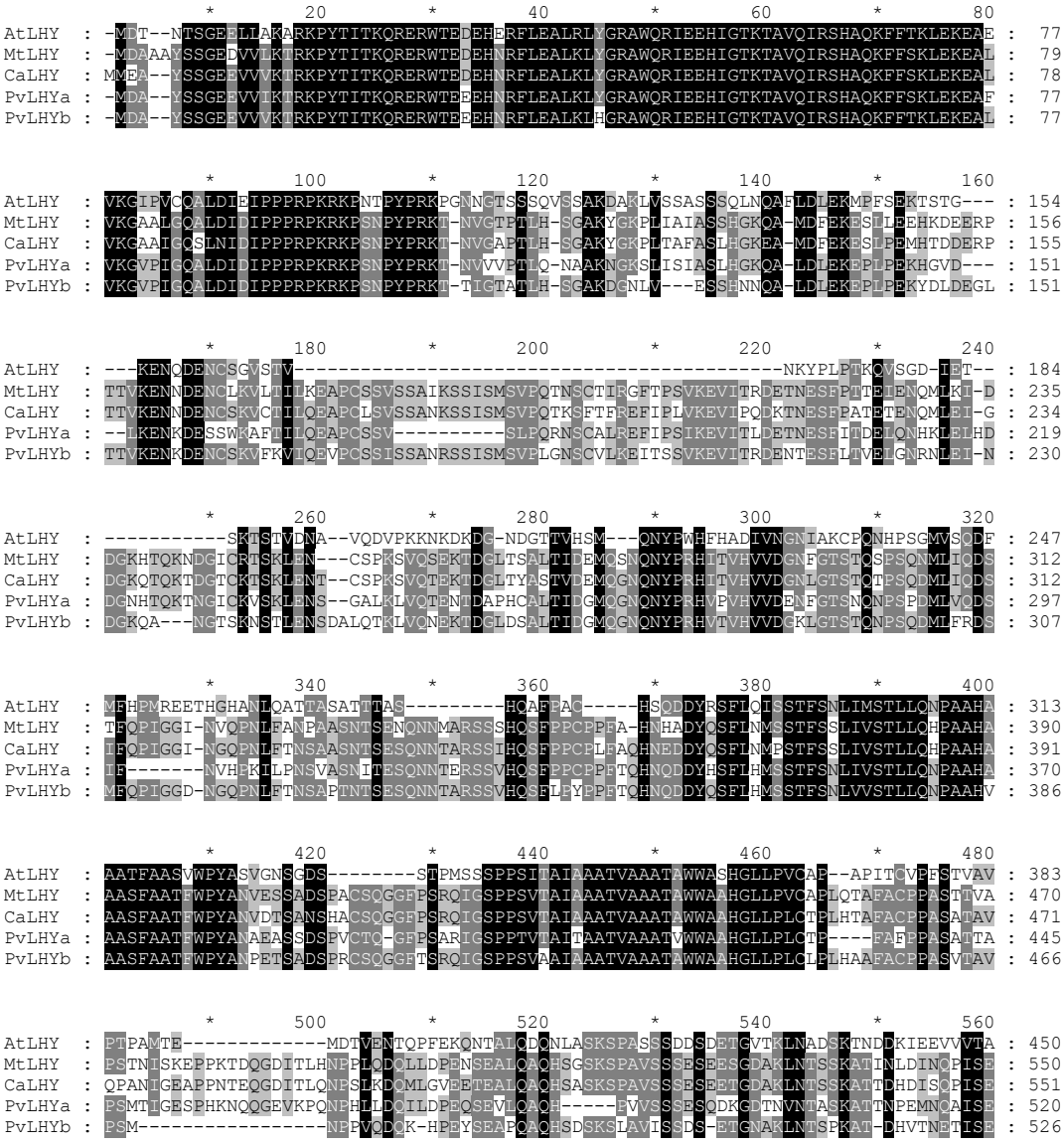


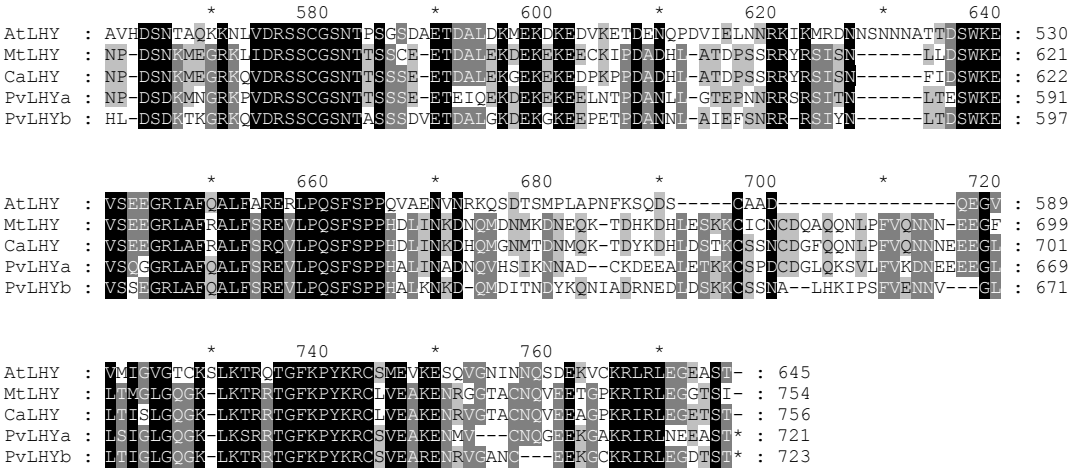

AtGI : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEDYASGVR---LPRNWMHLHFLRAIGLAMSMRAG : 383
MtGI : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEEYASGIR---LPRNWMHLHFLRAIGTAMSMRAG : 388
CaGI : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEEYASGIR---LPRNWMHLHFLRAIGTAMSMRAG : 388
PsGI : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEEYASGIR---LPRNWMHLHFLRAIGTAMSMRAG : 386
PvGIa : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEDYASGIR---LPRNWMHLHFLRAIGTAMSMRAG : 390
PvGIb : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEEYACGIR---LPRNWMHFHFLRAIGTAMSMRSG : 387
GmGIa : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEDYASGIRAKNLPRNWMHLHFLRAIGTAMSMRAG : 391
GmGIb : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEDYASGIR---LPRNWMHLHFLRAIGTAMSMRAG : 388
GmGIc : LFHRYYAIATPSATQRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEDYACGIR---LPRNWMHFHFLRAIGTAMSMRSG : 391
AtGI : VAADAAAALLFRILSQPALLFPPISQVEGVEIQHAPIGGYSSNYRKQIEVPAAEATIEATAQGIASMLCAHGPEVEWRIC : 463
MtGI : IAADAAAALLFRILSQPALLFPPLRQVDGVEVQHEPLGGYISSYSKQIEVPSAEASIDATAQGIASMLCAHGPEVEWRIC : 468
CaGI : IAADAAAALLFRILSQPALLFPPLRQVDGVEVQHEPLGGYISSYSKQIEVPAAEASIDATAQGIASLLCAHGPEVEWRIC : 468
PsGI : IAADAAAALLFRILSQPALLFPPLRQVDGVEVQHEPLGGYISSYSKQIEVPAAEASIDATAQGIASMLCAHGPEVEWRIC : 466
PvGIa : IAADAAAALLFRILSQPALLFPPLRHVDGVEVQHEPLGGYISSYKKQIEVPAAEASIEATAQGIASMLCAHGPEVEWRIC : 470
PvGIb : VAADAAAALLFRILSQPALLFPPLRQVDGVEIQHEPLGGYISSNRKQIEAAAAEATIEATAQGIASMLCAHGPEVEWRIC : 467
GmGIa : IAADAAAALLFRILSQPALLFPPLRQVDGVEVQHEPLGGYISSYKKQIEVPAAEASIEATAQGIASMLCAHGPEVEWRIC : 471
GmGIb : IAADAAAALLFRILSQPALLFPPLRQVDGVEVQHEPLGGYISSYKKQIEVPAAEASIEATAQGIASMLCAHGPEVEWRIC : 468
GmGIc : VAADAAAALLFRILSQPALLFPPLRQVDGVEVQHEPLGGYISSNRKQIEAASAEATIEATAQGIASVLCAHGPEVEWRIC : 471
AtGI : TIWEAAYGLIPLNSSAVDLPEIIVATPLQPPILSWNLYIPLLKVLEYLPRGSPSEACLMKIFVATVETILSRTFPPESSR : 543
MtGI : TIWEAAYGLIPASSSAVDLPEIIVAAPLQPPILSWNLYIPLLKVLEYLPRGSPSEACLMKIFAATVEAILQRTFPPESSR : 548
CaGI : TIWEAAYGLIPASSSAVDLPEIIVASPLQPPTLSWNLYIPLLKVLEYLPRGSPSEACLMKIFAATVEAILQRTFPPESSR : 548
PsGI : TIWEAAYGLIPANSSAVDLPEIIVAAPLQPPILSWNLYIPLLKVLEYLPRGSPSEACLMKIFAATVEAILQRTFPPESSR : 546
PvGIa : TIWEAAYGLIPTSSSAVDLPEIIVATPLQPPILSWNLYIPLLKVLEYLPRGSPSEACLMKIFAATVEAILQRTFPPESTR : 550
PvGIb : TIWEAAYGLIPLNSSAVDLPEIVVATPLQPPILSWNLYIPLLKVLEYLPRGSPSEACLMKIFVATVEAILQRTFPAEPIS : 547
GmGIa : TIWEAAYGLIPTSSSAVDLPEIIVATPLQPPILSWNLYIPLLKVLEYLPRGSPSEACLMKIFAATVEAILQRTFPPESTR : 551
GmGIb : TIWEAAYGLIPTSSSAVDLPEIIVATPLQPPVLSWNLYIPLLKVLEYLPRGSPSEACLMKIFAATVEAILQRTFPPESTR : 548
GmGIc : TIWEAAYGLIPLNSSAVDLPEIVVATPLQPPILSWNLYIPLLKVLEYLPRGSPSEACLMKIFVATVEAILQRTFPAESTS : 551
AtGI : ELTRKARSSFTTR--SATKNLAMSELRAMVHALFLESCAGVELASRLLFVVLTVCVSHEAQSSGSKRPRSEYASTTENIE : 621
MtGI : EQNRKANYLFGLG--SASKNLAVAELRTMVHSLFLESCASVELSSRLLFVVLTVCVSHEAQFSGSKKPRGEDNYSVEEII : 626
CaGI : EHNRKANYLFGIG--SASKNLAVAELRTMVHSLFLESCASVELASRLLFVVLTVCVSHEAQFSGSKKPRGEDNYSVEEII : 626
PsGI : EQNRKASYLFGIG--SASKNLAVAELRTMVHSLFLESCASVELSSRLLFVVLTVCVSHEAQFSGSKKPRGEDNYSVEEII : 624
PvGIa : EQSRKSKYLSIIG--SASKNLAVAELRTMVHSLFLESCASVELASRLLFVVLTVCVSHEAQFSGSKRPRGEDNYPAEEII : 628
PvGIb : DQKRKTRYYFVG---SASKNLAVAELRTMVHSLFLESCASVELASRLLFVVLTVCVSHEVKFNGSKKPRGEDNYL-EEVI : 623
GmGIa : EQNRKSKYLAGIG--SASKNLAMAELRTMVHSLFLESCASVELASRLLFVVLTVCVSHEAQFSGSKRPRGEDNYSAEDII : 629
GmGIb : EQNRKSKYLAGIGFGSASKNLAVAELRTMVHSLFLESCASVELASRLLFVVLTVCVSHEAQFSGSKRPRGEDNYSSEDII : 628
GmGIc : DHKIKTRYYCVG---SASKNLAVAELRTMVHSLFLESCASVELASRLLFVVLTVCVSHEVQFNGSKKPRGENNYLVEEII : 628
AtGI : ANQPVSNNQTANRKSRNVKGQGPVAAFDSYVLAAVCALACEVQLYPMISGGGNFSNSAVAGTITKPVKINGS-------- : 693
MtGI : EDLQAISEIRKERKNRKVKKQGPVAAFDSYVMAAVCALACELQLFPLMSRGNNHSVSNNVQDIAKPVTLHGS-------- : 698
CaGI : DDLQAISESRKERKNRKVKKQGPVAAFDSYVMAAVCALACELQLFPLMSRGNNHSVSNNVQDIAKPVTLNGS-------- : 698
PsGI : EDLQAISEIRKERKNRKVKKQGPVAAFDSYVMAAVCALACELQLFPLISRGNNHSLSNNGQDIAKPVTLHGS-------- : 696
PvGIa : EDLQTS-ENQKESKNRKMKKQGPVAAFDSYVLAAVCALACELQLFPLISRGSNNLVSNNAQVIAKPAKLNGSSHRQNGSS : 707
PvGIb : EDIFPVSESQKETKNRKIKKQGPVSAFDSYVLAAVCALACELQLFPTVSRGNNRSVLNKVEKIAKPSKIDES-------- : 695
GmGIa : EDLQTS-ENQKVSKNRKLKKQGPVAAFDSYVLAAVCALACELQLFPLISCGNNRLASNNVQDIAKPVRLNGS-------- : 700
GmGIb : EDLQTS-ENQKESKNRKLKKQGPVAAFDSYVLAAVCALACELQLFPLISRGNNHLASNKVQDIAKPVRLNGS-------- : 699
GmGIc : EDLQAVSESQKETKNRKMKKQGPVAAFDSYVLAAVCALACELQLFPMISRGDNHSVPNNVQNIAKPVKIDGS-------- : 700
AtGI : ---------------SKEYGAGIDSAISHTRRILAILEALFSLKPSSVGTPWSYSSSEIVAAAMVAAHISELFRRSKALT : 758
MtGI : ---------------SQDTQNGLDSAVRHTHRILAILEALFSLKPSSVGTPWSYSSNEIVAAAMVAAHVSELFRRSKACM : 763
CaGI : ---------------SQDTQNGIDSAVRHTHRILSILEALFSLKPSSVGTPWSCSSNEIVAAAMVAAHVSELFRRSKACM : 763
PsGI : ---------------SQDTQNGLESAVRHTHRILAILEALFSLKPSSVGTPWSYSSNEIVAAAMVAAHVSELFRRSKACM : 761
PvGIa : HIRQNGSSHIRQNGSSHDIRNGLDSAVRHTHRILAILEALFSLKPSSVGTPWSYSSNEIVAAAMVAAHVSELFRRSKACM : 787
PvGIb : ---------------SHALQKCIDSAIRHTHRILAILEALFSLKPSSVGTSWSYSSNEIVAAAMVAAHISELFRRSKACM : 760
GmGIa : ---------------SHELQNGLDSAMRHTHRILAILEALFSLKPSSVGTPWSYSSNEIVAAAMVAAHVSELFRRSKTCM : 765
GmGIb : ---------------SHEIRNGLDSAVRHTHRILAILEALFSLKPSSVGTPWSYSSNEIVAAAMVAAHVSELFRRSKACM : 764
GmGIc : ---------------SHALQNGIDSAIRHTHRILAILEALFSLKPSSVGTSWSYSSNEIVAAAMVAAHISELFRRSKACM : 765
AtGI : HALSGLMRCKWDKEIHKRASSLYNLIDVHSKVVASIVDKAEPLEAYLKNTPVCKDSVTCLNWKQENTCASTTCFDTAVTS : 838
MtGI : HALSVLIRCKWNKEIHSRASSLYNLIDIHSKVVASIVNKAEPLEATLIHAPIYKDALVCHDGKPKNRSENGSCSDPGQTS : 843
CaGI : HALSVLIRCKWDKEIHSRASSLYNLIDIHSKVVASIVNKAEPLEATLIHAPIYKDSLVCHDGKPKNRSENGSCSDPGQTS : 843
PsGI : HALSVLIRCKWNKEIHSRASSLYNLIDIHSKVVASIVNKAEPLEATLIHAPIYKDSLVCHDGKPKNRSENGGCSDSRQTS : 841
PvGIa : HALSVLIRCKWDNEIHSRASSLYNLIDIHSKAVASIVNKAEPLEATLIHAPIWRDSRIYCGNKKQNQCESN-CFDPGQTS : 866
PvGIb : RALSVLIRCKWDDEIHSRASSLYNLIDIHSKAVASIVNKAEPFGATLIHTPICKDP--CGSIKGQNQFENSSHLLPGQTS : 838
GmGIa : HALSVLIRCKWDNEIHSRASSLYNLIDIHSKAVASIVNKAEPLEATLIHVPIWKDSLVCVGVKKQNQCESSSCFAPGQTS : 845
GmGIb : HALSVLIRCKWDNEIHSRASSLYNLIDIHSKAVASIVNKAEPLEATLIHAPIRKDSLVCVGVKKQNQCESSSCFDAGRTS : 844
GmGIc : RALSVLMRYKWDDEIHSRASSLYTLIDIHRKAVASIVNKAEPLGATFIHTPICKDP--CDRSKKKNQCENSRCLDPGQTS : 843


AtGI : ASR-------TEMNPRGNHKYARHSD-EGSGRPSEKGIKDFLLDASDLANFLTADRLAGFYCGTQKLLRSVLAEKPELSF : 910
MtGI : IVP---SADSTPSKHIHKSGRTPCSNEEASGYNLGKGVISFSLEASDLANFLTMDRHIGLNCNTQIFLISMLSEKQELCF : 920
CaGI : IVPLEPSEDSTPSKHSHKSGRTPCSNEAASGYNMGKGVIGFSLDASDLANFLTMDRHIGLNCNTQIFLIPMLAEKQELCF : 923
PsGI : TVP---SEDSTPSKHSHKSGRTPCSNEEASGYNLGKGVIGFSLEASDLANFLTMDRHIGLNCNTQIFLISMLSEKQELCF : 918
PvGIa : IIP---SADSFPSKPVHTSKKTPCSN-EAAGCTLGKGVSGFPLDASDLANFLTVDRHIGLNCNAQIFLRSMLAEKQELCF : 942
PvGIb : TSI---STDSSPLELGKRCDRTSYSN-EASGCTFGKGATSLAFDASDLANFLTMDRHIGFNCSAQIFLRSRLAEKQELCF : 914
GmGIa : VVP---SEDSFPSKVDHNSQKTPCSK-DASDYTLGKGVIGFSLDASDLANFLTMDRHIGLNCNGQIFLRSTLAEKQELCF : 921
GmGIb : VVP---SEDSFPSKLDHNSNKTPCPK-GASDYTLGKGVIGFSLDASDLANFLTMDRHIGLNCNGQIFLRSTLAEKQELCF : 920
GmGIc : TST---SEDSSRSKFSKKSERTSYSN-EASGCTFEKATTSLPFDASDLANFLTVDRHIGFNCSAQIFLRSRLAEKQQLCF : 919
AtGI : SVVSLLWHKLIAAPEIQPTAESTSAQQGWRQVVDALCNVVSATEAKAAAAVVLQAERELQPWIAKDDEEGQKMWKINQRI : 990
MtGI : SVVSLLWHKLIASPETQPCSESTSAQQGWRQVVDALCNVVSASEAKAAIAVVLQAEKELQPWIAKDDDLGQKMWRINQRI : 1000
CaGI : SVVSLLWHKLIASPETQPCSESTSAQQGWRQVVDALCNVVSASEAKAAIAVVLQAEKELQPWIAKDDDLGQKMWRVNQRI : 1003
PsGI : SVVSLLWHKLIASPETQPCSESTSAQQGWRQVVDALCNVVSAAEAKAAIAVVLQAEKELQPWIAKDDDLGQKMWRINQRI : 998
PvGIa : SVVSLLWHKLIASPETQPCAESTSAQQGWRQVVDALCNVVSASEPTKAAIAVVLQAERELQPWIAKDDDLGQKMWRINQRI : 1022
PvGIb : SVVSLLWQKLIASPETQPCAESTSAQQGWRQVVDALCQVVSASEMKAAIAVVLQAEKELQPWIAKDDNLGQKMWRINQRI : 994
GmGIa : SVVSLLWHKLIASPETQPCAESTSAQQGWRQVVDALCNVVSASEPTKAAIAVVLQAERELQPWIAKDDDLGQKMWRINQRI : 1001
GmGIb : SVVSLLWHKLIASPETQPCAESTSAQQGWRQVVDALCNVVSASEPTKAAIAVVLQAERELQPWIAKDDLSGQKMWRINQRI : 1000
GmGIc : SVVSLLWQKLIGSPETQPCAESTSAQQGWRQVADALCKVVSASEPTKAAMAVVLQAEKELQPWIAKDDNLGQKLWRINQRI : 999
AtGI : VKVLVELMRNHDREESLVILASASDLLLRATDGMLVDGEACTLPQLELLEATARAIQPVLAWGPSGLAVVDGLSNLLKCR : 1070
MtGI : VKLIVELMRNHDSAESLVILASASDLLLRATDGMLVDGEACTLPQLELLEATARAIQPVLEFGEPGLAVADGLSNLLKCR : 1080
CaGI : VKLIVELMRNHDSSESLVILASASDLLLRATDGMLVDGEACTLPQLELLEATARAVQPVLEFGESGLAVADGLSNLLKCR : 1083
PsGI : VKLIVELMRNHDSSESLVILASASDLLLRATDGMLVDGEACTLPQLELLEATARAIQPVLEFGEPGMAVADGLSNLLKCR : 1078
PvGIa : VKLIVELMRNDESAESLVIVASSSDLLLRATDGMLVDGEACTLPQLELLEATARAVQPVLEFGESGLAVADGLSNLLKCR : 1102
PvGIb : VKLIIELMRNHDSLESLVVVASASDLLLRATDGMLVDGEACTLTQLKLLEATAKAVQPVIELGESGFAVADGLSNLLKCR : 1074
GmGIa : VKLIVELMRNHEISESLVIVASSSDLLLRATDGMLVDGEACTLPQLELLEATARAVQPVLEFGESGLAVADGLSNLLKCR : 1081
GmGIb : VKLIVELMRNHETAESLVIVASSSDLLLRATDGMLVDGEACTLPQLELLEATARAVQPVLEFGESGLAVADGLSNLLKCR : 1080
GmGIc : VKLITELLRNHDNIESLVIVASASDLLLRATDGMLVDGEACTLPQLKLLEATAKAVQPVIELGESGLGVADGLSNLLK-- : 1077
AtGI : LPATIRCLSHPSAHVRALSTSVLRDIMNQSSIPIKVTPKLPTTEKNGMNSPSYRFFNAASIDWKADIQNCLNWEAHSLLS : 1150
MtGI : LAATIRCLCHPSAHVRALSVSVLRDILHTGSIR--CSPKPLRI--NGSHNPSYQYFKLDVVDWQADIEKCMAWEAHSRIS : 1156
CaGI : LAATIRCLSHPSAHVRTLSVSVLRDILHTSSIR--CNPKPLRI--NGNHNPSYQYFKLDVVDWQTDIEKCLTCEAHSRIS : 1159
PsGI : LAATIRCLCHPSAHVRTLSVSVLRDILHTGSIR--CSPKPLRI--NGNHNPSYPYFKLDVVDWQADIEKCLTCEAHSRIS : 1154
PvGIa : LSATIRCLSHPSAHVRALSISVLRDILHTGSIR--YNLKPRRI--NGTHNPSYQYFNSDAIDWQADIEKCLTWEAHSRLS : 1178
PvGIb : LPATIRCLSHPSAHVRALSTSVLRDILHTCSIRR-YSPKRPQK--NDIHN---QYFNLDVIDWQADINKCLTWEAHSQLL : 1148
GmGIa : LSATIRCLSHPSAHVRALSISVLRDILHTGSIR--CSPKPRRL--NGTHNPSYQYFNLDVIDWQADIEKCLTWEAHSRLS : 1157
GmGIb : LSATIRCLSHPSAHVRALSISVLRDILHTGSIR--CSPKPRRL--NGTHNPSYQYFNLLAVDWQADIEKCLTWEAHSRLS : 1156
GmGIc : --------------VNKLSSFELEPILNVKF*---------- : 1094
AtGI : TTVPTQFLDTAARELGCTISLSQ : 1173
MtGI : AGLPIKFLDTAAKELGCAISV-- : 1177
CaGI : SGLPIKFLDTAAKELGCAISI-- : 1180
PsGI : AGLPIKFLDTAAKELGCAISI-- : 1175
PvGIa : TRLPINFLDTAAKELGCNISM*- : 1199
PvGIb : NGVSIEYLNTAAFDLGFGISI*- : 1169
GmGIa : NGLSINFLDTAAKELGCTISM*- : 1178
GmGIb : NGLSINFLDTAAKELGCTISM*- : 1177
GmGIc : ---------- : -

Appendix 3.7 Identity matrix (%) derived from the multiple sequence alignment of LATE ELONGATED HYPOCOTYL (LHY) genes in the species *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt), *Cicer arietinum* (Ca) and *Phaseolus vulgaris* (Pv). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

	Accession	AtLHY	MtLHY	CaLHY	PvLHYa	PvLHYb
AtLHY	AT1G01060		41.9	42.5	41.6	45.4
MtLHY	Medtr7g118330	41.9		84.0	67.3	68.9
CaLHY	LOC101500635	42.5	84.0		66.7	69.7
PvLHYa	Phvul.009G259650	41.6	67.3	66.7		64.1
PvLHYb	Phvul.010G120401	45.4	68.9	69.7	64.1	




AtLHY : AVHDSNTAQKKNLVDRSSCGSNTPSGSDAETDALDKMEKDKEDVKETDENQPDVIELNNRKIKMRDNNSNNNATTDSWKE : 530
MtLHY : NP-DSNKMEGRKLIDRSSCGSNTTSSCE-ETDALEKDEKEKEECKIPDADHI-ATDPSSRRYRSISN------LLDSWKE : 621
CaLHY : NP-DSNKMEGRKQVDRSSCGSNTTSSSE-ETDALEKGEKEKEDPKPPDADHI-ATDPSSRRYRSISN------FIDSWKE : 622
PvLHYa : NP-DSDKMNGRKPVDRSSCGSNTTSSSE-ETEIQEKDEKEKEELNTPDANLI-GTEPNNRRSRSITN------LTESWKE : 591
PvLHYb : HL-DSDKTKGRKQVDRSSCGSNTASSSDVETDALGKDEKGKEEPETPDANNI-AIEFSNRR-RSIYN------LTDSWKE : 597
AtLHY : VSEEGRIAFQALFARERLPQSFSPPQVAENVNRKQSDTSMPLAPNFKSQDS-----CAAD----------------QEGV : 589
MtLHY : VSEEGRLAFRALFSREVLPQSFSPPHDLINKDNQMDNMKDNEQK-TDHKDHLESKKCICNCDQAQQNLPFVQNNN-EEGF : 699
CaLHY : VSEEGRLAFRALFSRQVLPQSFSPPHDLINKDHQMGNMTDNMQK-TDYKDHLDSTKCSSNCDGFQQNLPFVQNNNEEEGL : 701
PvLHYa : VSCGGRLAFQALFSREVLPQSFSPPHALINADNQVHSIKNNAD--CKDEEALETKKCSPDCDGLQKSVLFVKDNEEEEGL : 669
PvLHYb : VSSEGRLAFQALFSREVLPQSFSPPHALKNKD-QMDITNDYKQNIADRNEDLDSKKCSSNA--LHKIPSFVENNV---GL : 671
AtLHY : VMIGVGTCKSLKTRQTGFKPYKRCSMEVKESQVGNINNQSDEKVCKRLRLECEASST- : 645
MtLHY : LTMGLGQGK-LKTRRTGFKPYKRCLVEAKENRGCTACNQVEETGPKRIRLEGGTSI- : 754
CaLHY : LTISLGQGK-LKTRRTGFKPYKRCLVEAKENRVGTACNQVEEAGPKRIRLEGETST- : 756
PvLHYa : LSIGLGQGK-LKSRRTGFKPYKRCSVEAKENMV---CNQGEEKGAKRIRLNEEASST* : 721
PvLHYb : LTIGLGQGK-LKTRRTGFKPYKRCSVEARENRVGANC---EEKGCKRIRLECDISST* : 723

Appendix 3.8 Maximum parsimony tree derived from the alignment of genes involved in circadian clock control in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Pisum sativum* (Ps) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using ClustalW and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications.

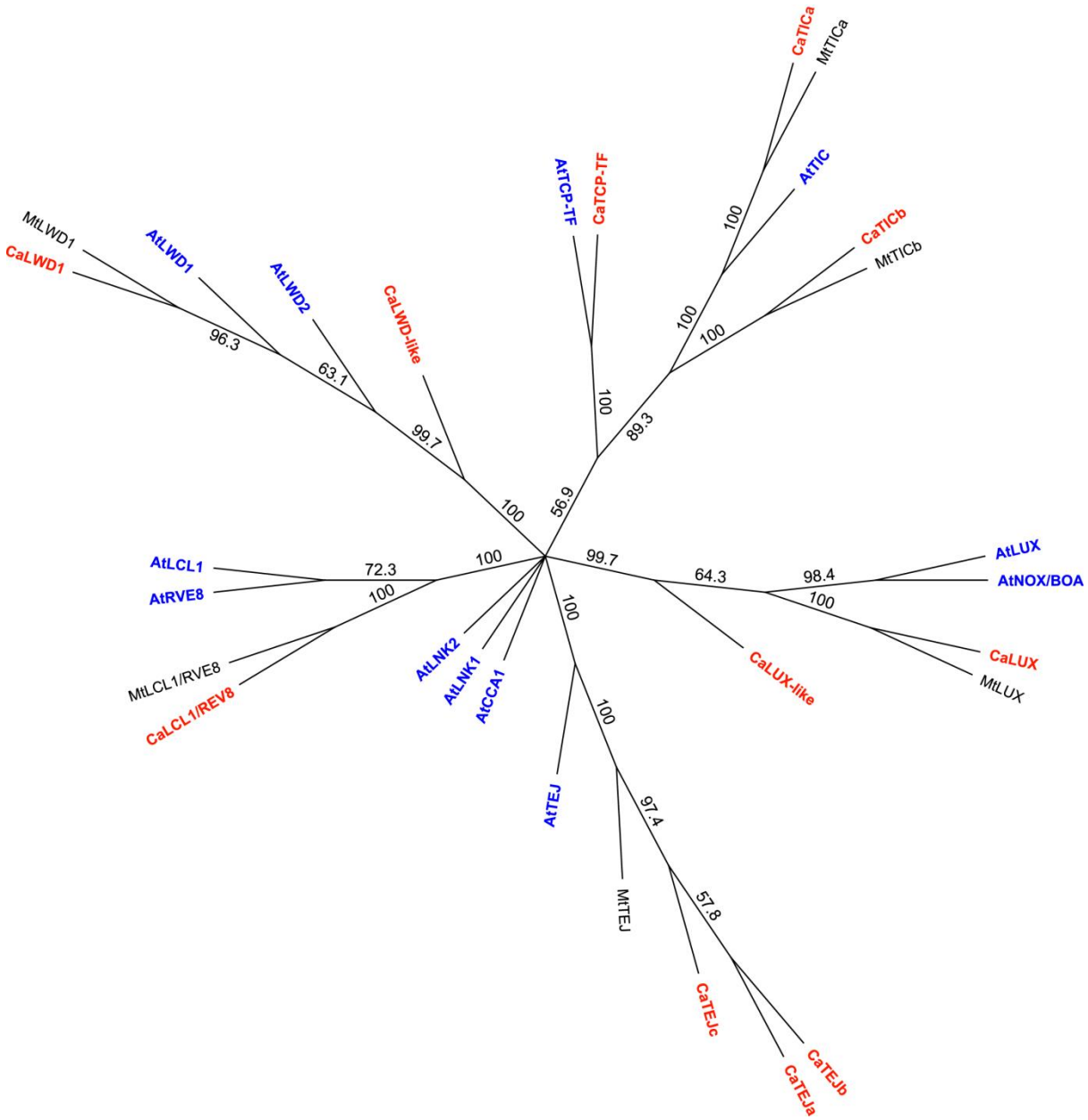


Table 1 Accession number of the sequences used in the alignment and tree construction of circadian clock genes in three plant species.

Arabidopsis thaliana	
LCL1	AT5G02840
RVE8	AT3G09600
TCP TF	AT5G08330
LWD1	AT1G12910
LWD2	AT3G26640
TIC	AT3G22380
TEJ	AT2G31870
LNK1	AT5G64170
LNK2	AT3G54500
CCA1	AT2G46830
LUX	AT3G46640
BOA	AT5G59570

Medicago truncatula	
LCL1/RVE8	Medtr1g067000
LWD1	Medtr7g084810
TICa	Medtr1g104710
TICb	Medtr1g072160
TEJ	Medtr3g029520
LUX	Medtr4g064730

Cicer arietinum	
LCL1/RVE8	LOC101514458
TCP TF	LOC101492981
LWD1	LOC101494250
LWD-like	LOC101515679
TICa	LOC101496754
TICb	LOC101504248
TEJa	LOC101507719
TEJb	LOC101509623
TEJc	LOC101507492
LUX	LOC101515043
LUX-like	LOC101494139

Appendix 3.9 Maximum parsimony tree derived from the alignment of *Cycling DOF factor* (CDF) genes in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Pisum sativum* (Ps) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using ClustalW and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications. The DOF family in *Arabidopsis* is formed by 16 genes (see Table 1). They were all included in the alignment and tree to ensure a correct identity of the chickpea sequences retrieved.

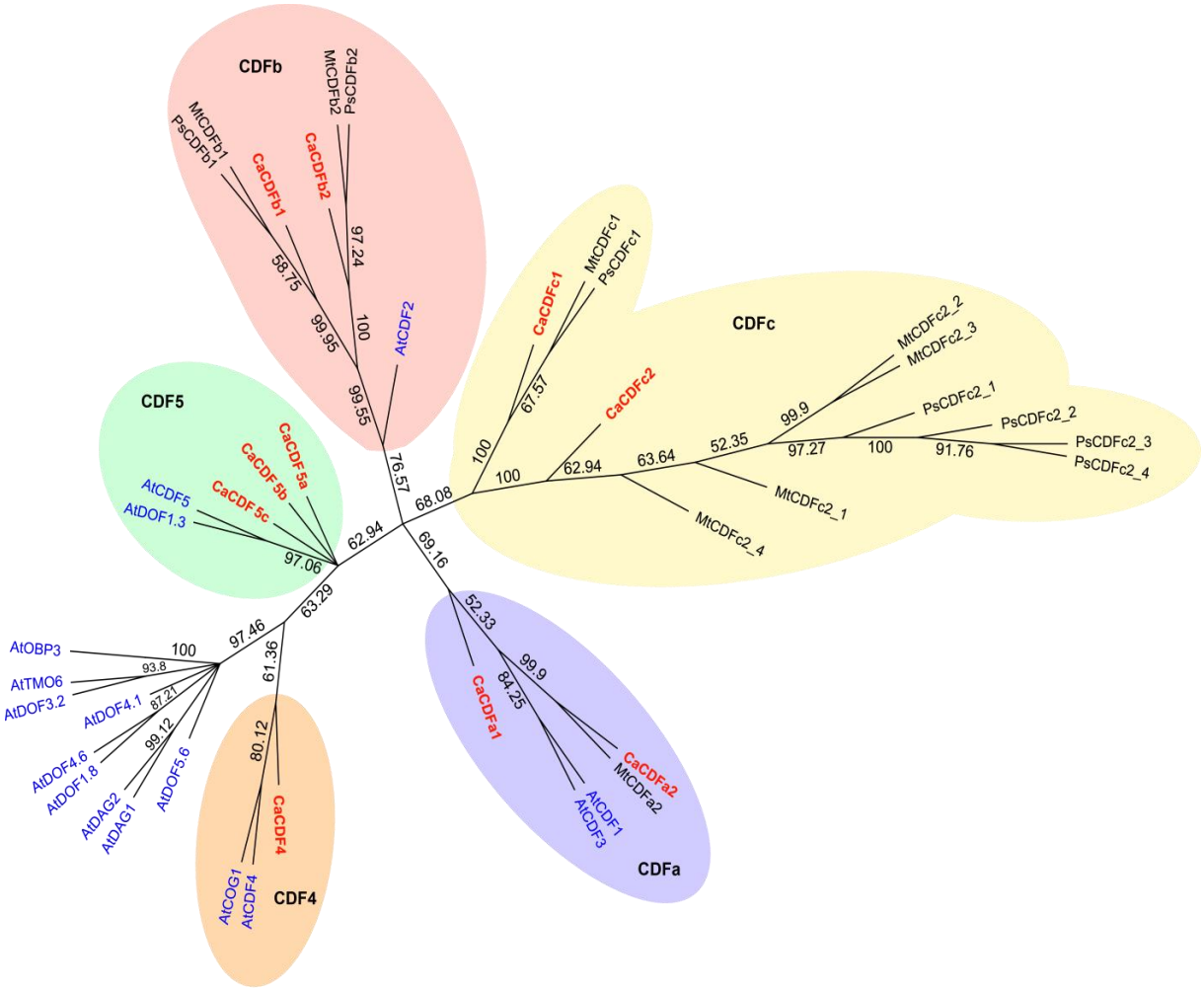


Table 1 Accession number of the sequences used in the alignment and tree construction of CDF genes in four plant species.

Arabidopsis thaliana	
CDF1	At5G62430
CDF3	At3G47500
CDF2	At5G39660
CDF4	At2G34140
CDF5	At1G69570
COG1	At1g29160
DAG1	At3g61850
DAG2	At2g46590
DOF1.3	At1g26790
DOF1.8	At1g64620
DOF3.2	At3g45610
DOF4.1	At4g00940
DOF4.6	At4g24060
DOF5.6	At5g62940
OBP3	At3g55370
TMO6	At5g60200
DOF1.8	At1g64620
DOF3.2	At3g45610
DOF4.1	At4g00940
DOF4.6	At4g24060
DOF5.6	At5g62940
OBP3	At3g55370
TMO6	At5g60200

Pisum sativum	
CDFb1	PsCam037510
CDFb2	PsCam048092
CDFc1	PsCam036807_SR
CDFc2_1	PsCam038247
CDFc2_2	PsCam038945
CDFc2_3	PsCam014419
CDFc2_4	PsCam014395

Medicago truncatula	
CDFa2	Medtr3g435480
CDFb1	Medtr7g010950
CDFb2	Medtr6g012450
CDFc1	Medtr4g082060
CDFc2_1	Medtr5g041420
CDFc2_2	Medtr5g041400
CDFc2_3	Medtr5g041380
CDFc2_4	Medtr5g041530

Cicer arietinum	
CDFa1	LOC101503333
CDFa2	LOC101496410
CDFb1	LOC101512379
CDFb2	LOC101490004
CDFc1	LOC101511051
CDFc2	LOC101503458
CDF4	LOC101502161
CDF45a	LOC101500722
CDF45b	LOC101489791
CDF45c	LOC101489693

Appendix 3.10 Maximum parsimony tree derived from the alignment of Pseudo-response regulator (PRR) and TIMING OF CAB EXPRESSION 1 (TOC1) genes in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using MUSCLE and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from1000 replications.

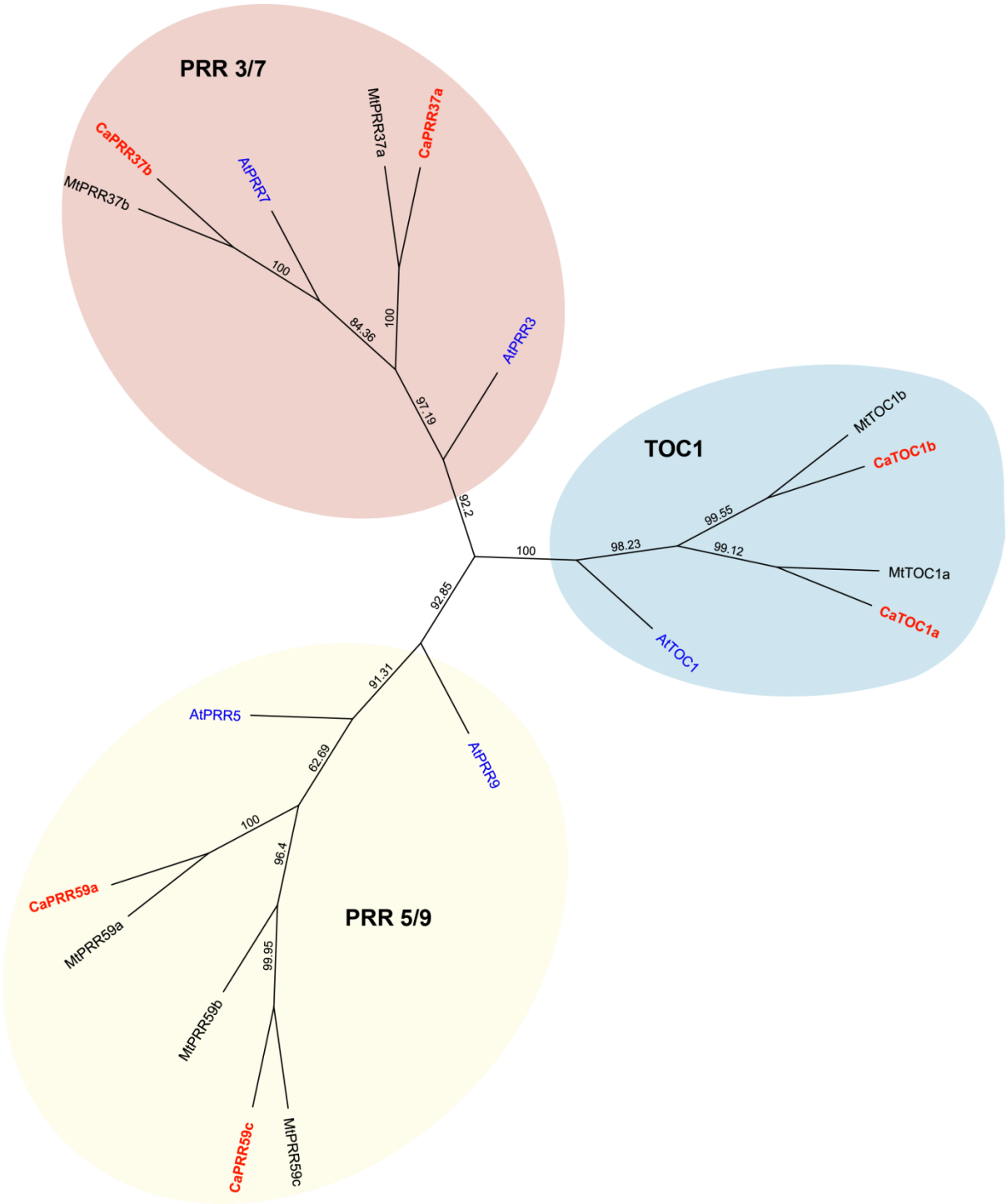


Table 1 Accession number of the protein sequences used in the alignment and tree construction of PRR and TOC1 genes in three plant species.

Medicago truncatula	
TOC1a	Medtr4g108880
TOC1b	Medtr3g037390
PRR37b	Medtr1g067110
PRR59a	Medtr3g092780
PRR37a	Medtr4g061360
PRR59b	Medtr8g024260
PRR59c	Medtr7g118260

Cicer arietinum	
TOC1a	LOC101500788
TOC1b	LOC101514145
PRR37b	LOC101512945
PRR59a	LOC101514597
PRR37a	LOC101494510
PRR59c	LOC101498831

Arabidopsis thaliana	
TOC1	AT5G61380
PRR3	AT5G60100
PRR5	AT5G24470
PRR7	AT5G02810
PRR9	AT2G46790

Appendix 3.11 Maximum parsimony tree derived from the alignment of seven genes involved in the flowering autonomous pathway in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using ClustalW and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications.

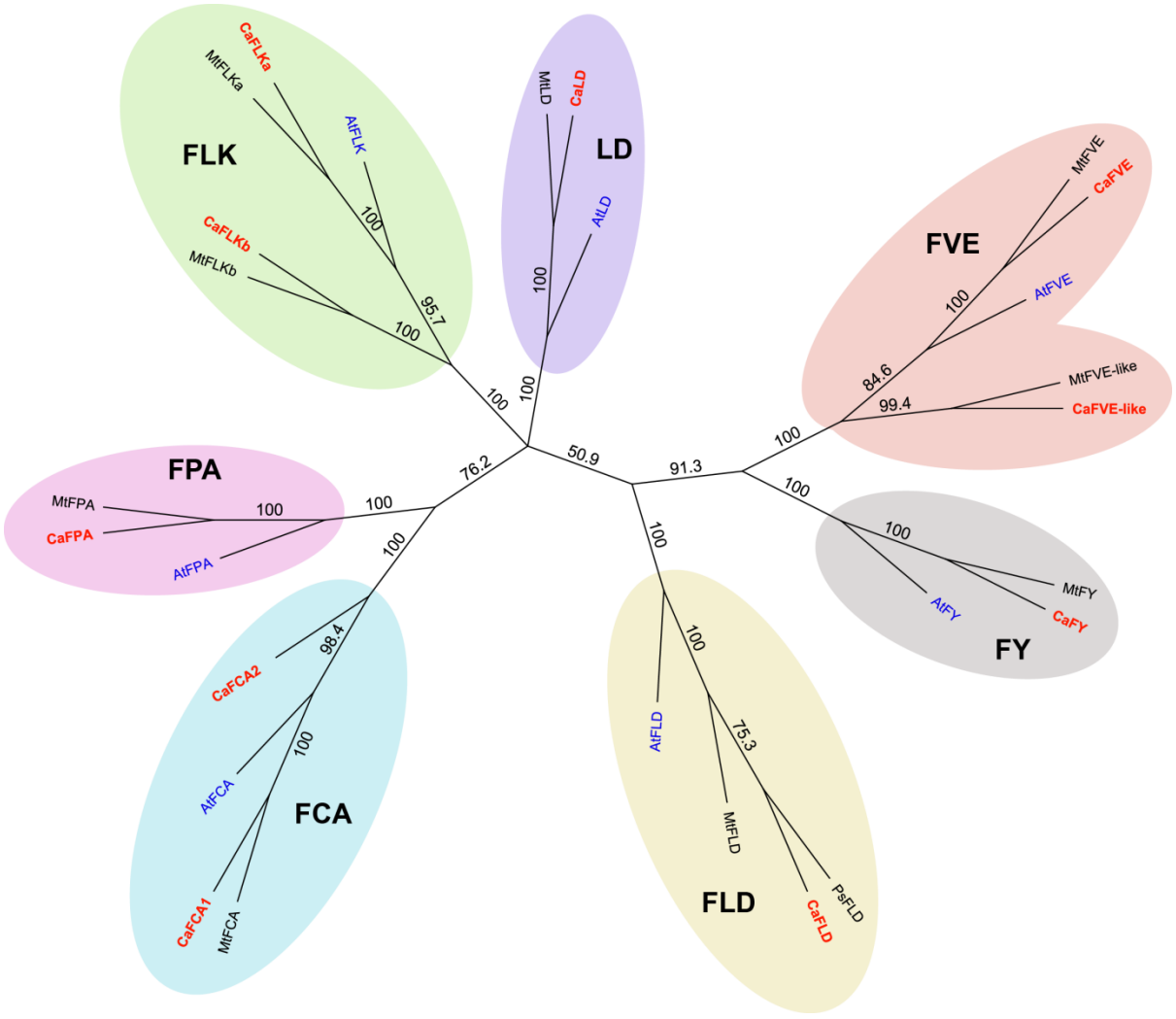


Table 1 Accession number of the protein sequences used in the alignment and tree construction.

Cicer arietinum	
CaFCA1	LOC101498212
CaFCA2	LOC101495434
CaFY	LOC101514205
CaFLD	LOC101490188
CaFVE	LOC101501320
CaFVE-like	LOC101509771
CaFPA	LOC101500681
CaFLKa	LOC101506610
CaFLKb	LOC101509502
CaLD	LOC101505423

Pisum sativum	
PsFLD	PsCam001633

Medicago truncatula	
MtFCA	Medtr4g122650
MtFY	Medtr8g038550
MtFLD	Medtr1g050535
MtFVE	Medtr2g039250
MtFVE-like	Medtr2g100090
MtFPA	Medtr4g068120
MtFLKa	Medtr7g115340
MtFLKb	Medtr1g070380
MtLD	Medtr7g108390

Arabidopsis thaliana	
AtFCA	AT4G16280
AtFY	AT5G13480
AtFLD	AT3G10390
AtFVE	AT2G19520
AtFPA	AT2G43410
AtFLK	AT3G04610
AtLD	AT4G02560

Appendix 3.12 Maximum parsimony tree derived from the alignment of *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (SPL) gene family in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Pisum sativum* (Ps) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions in Table 1 were aligned using MAFFT and PAUP* software was used to build the tree. Numbers in branches represent bootstrap consensus support from 1000 replications.

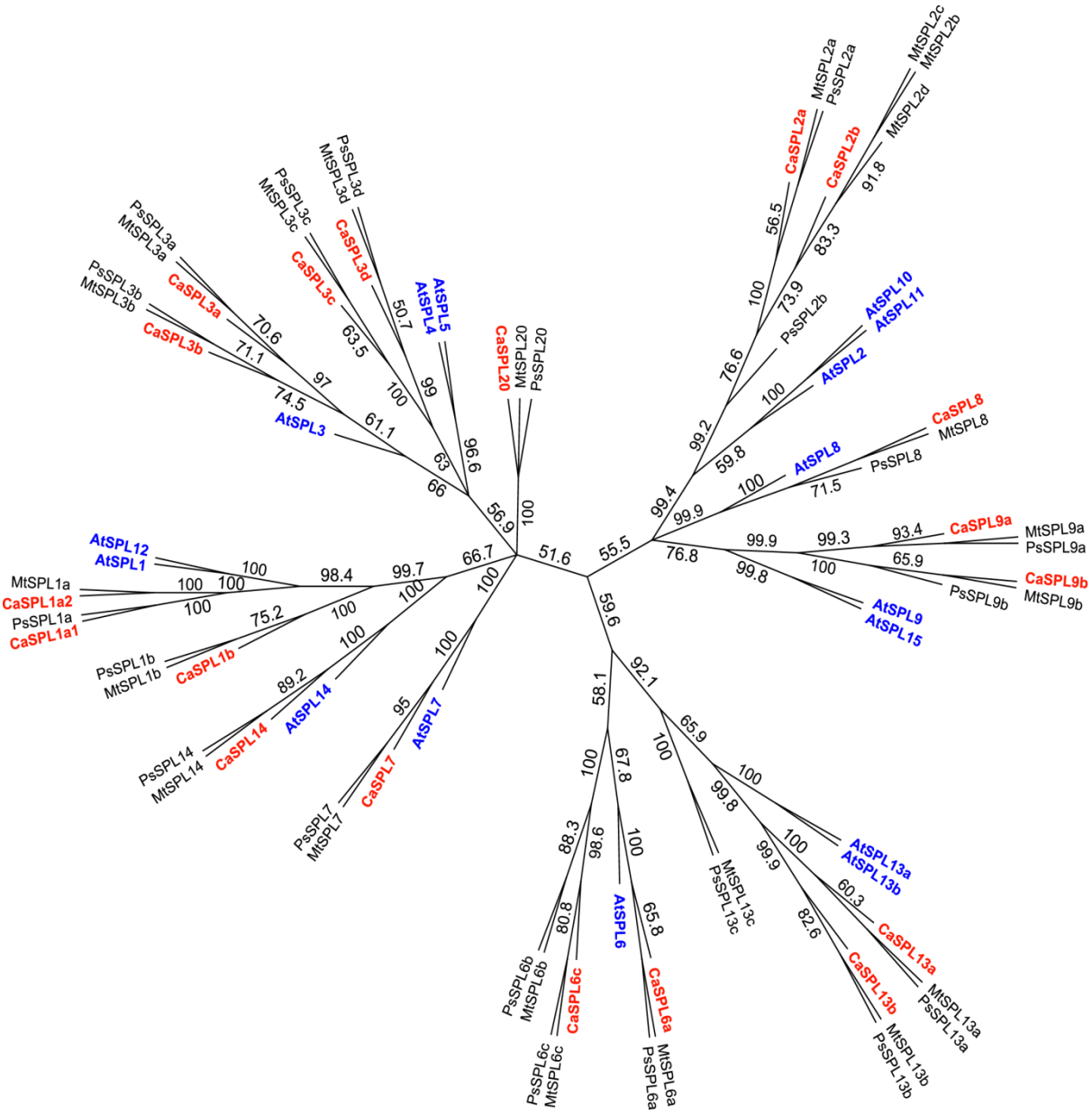
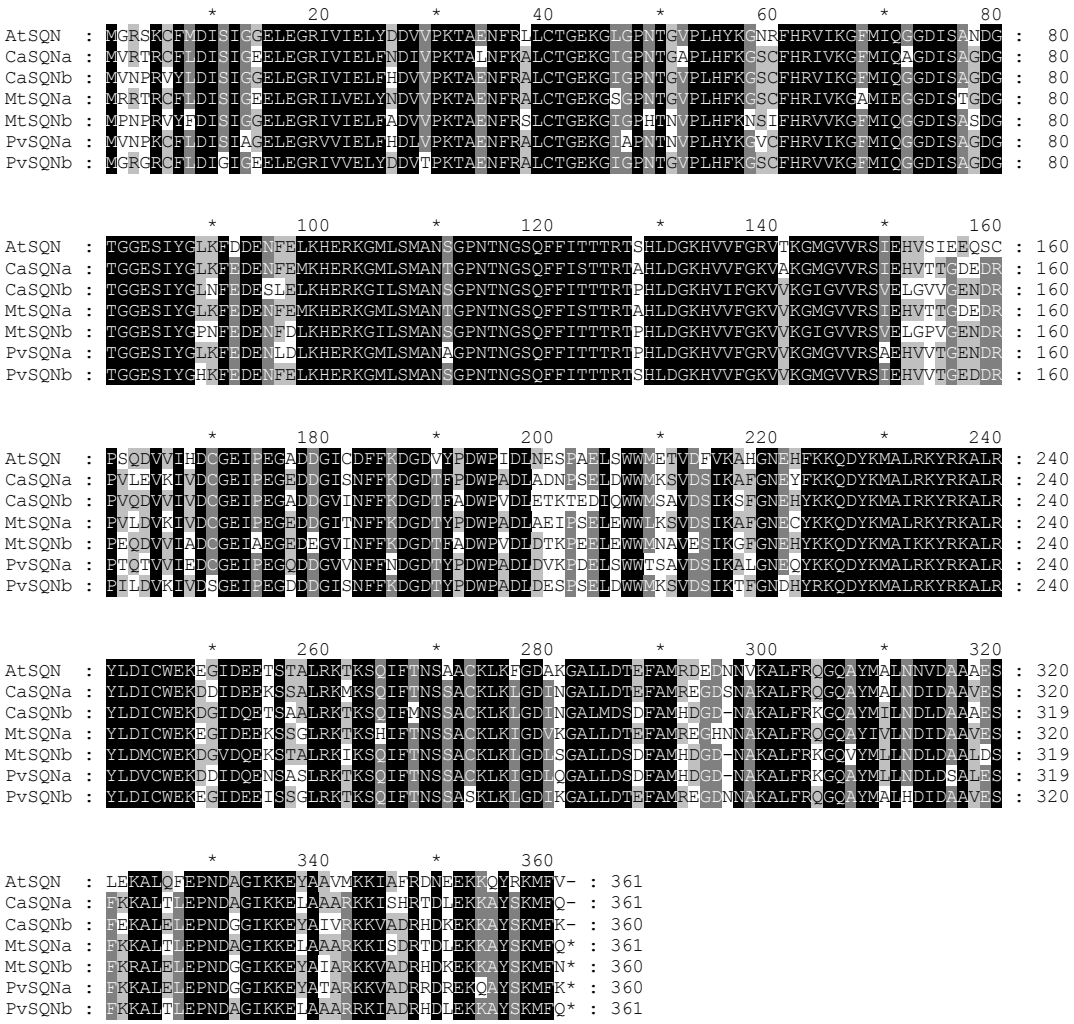


Table 1 Accession number of the sequences used in the alignment and tree construction of SPL genes in four plant species.

Arabidopsis thaliana		***Cicer arietinum***		***Pisum sativum***		***Medicago truncatula***	
SPL1	AT2G47070	**SPL1a1**	LOC101492418	**SPL1A**	PsCam048790	**SPL1A**	Medtr7g110320
SPL2	AT5G43270	**SPL1a2**	LOC101512064	**SPL1B**	PsCam048218	**SPL1B**	Medtr2g046550
SPL3	AT2G33810	**SPL1b**	LOC101503890	**SPL2A**	PsCam037189	**SPL2A**	Medtr3g085180
SPL4	AT1G53160	**SPL2a**	LOC101505236	**SPL2B**	PsCam037577	**SPL2B**	Medtr8g080670
SPL5	AT3G15270	**SPL2b**	LOC101507206	**SPL3A**	PsCam033864	**SPL2C**	Medtr8g080680
SPL6	AT1G69170	**SPL3a**	LOC101501522	**SPL3B**	PsCam017682	**SPL2D**	Medtr8g080690
SPL7	AT5G18830	**SPL3b**	LOC101512968	**SPL3C**	PsCam038909	**SPL3A**	Medtr2g014200
SPL8	AT1G02065	**SPL3c**	LOC101497340	**SPL3D**	PsCam047017	**SPL3B**	Medtr4g088555
SPL9	AT2G42200	**SPL3d**	LOC101498218	**SPL6A**	PsCam033315	**SPL3C**	Medtr8g463140
SPL10	AT1G27370	**SPL6a**	LOC101505285	**SPL6B**	PsCam036737	**SPL3D**	Medtr2g078770
SPL11	AT1G27360	**SPL6c**	LOC101511918	**SPL6C**	PsCam025564	**SPL6A**	Medtr5g046670
SPL12	AT3G60030	**SPL7**	LOC101502900	**SPL7**	PsCam057199	**SPL6B**	Medtr4g109770
SPL13A	AT5G50570	**SPL8**	LOC101501438	**SPL8**	PsCam001061	**SPL6C**	Medtr2g461920
SPL13B	AT5G50670	**SPL9a**	LOC101492646	**SPL9A**	PsCam039166	**SPL7**	Medtr2g020620
SPL14	AT1G20980	**SPL9b**	LOC101502532	**SPL9B**	PsCam037965	**SPL8**	Medtr8g005960
SPL15	AT3G57920	**SPL13a**	LOC101507093	**SPL13A**	PsCam046016	**SPL9A**	Medtr1g053715
		SPL13b	LOC101507938	**SPL13B**	PsCam055995	**SPL9B**	Medtr7g092930
		SPL14	LOC105851070	**SPL13C**	PsCam012994	**SPL13A**	Medtr3g099080
		SPL20	LOC101496293	**SPL14**	PsCam048786	**SPL13B**	Medtr8g096780
				SPL20	PsCam037459	**SPL13C**	Medtr7g028740
						SPL14	Medtr1g035010
						SPL20	Medtr7g444860

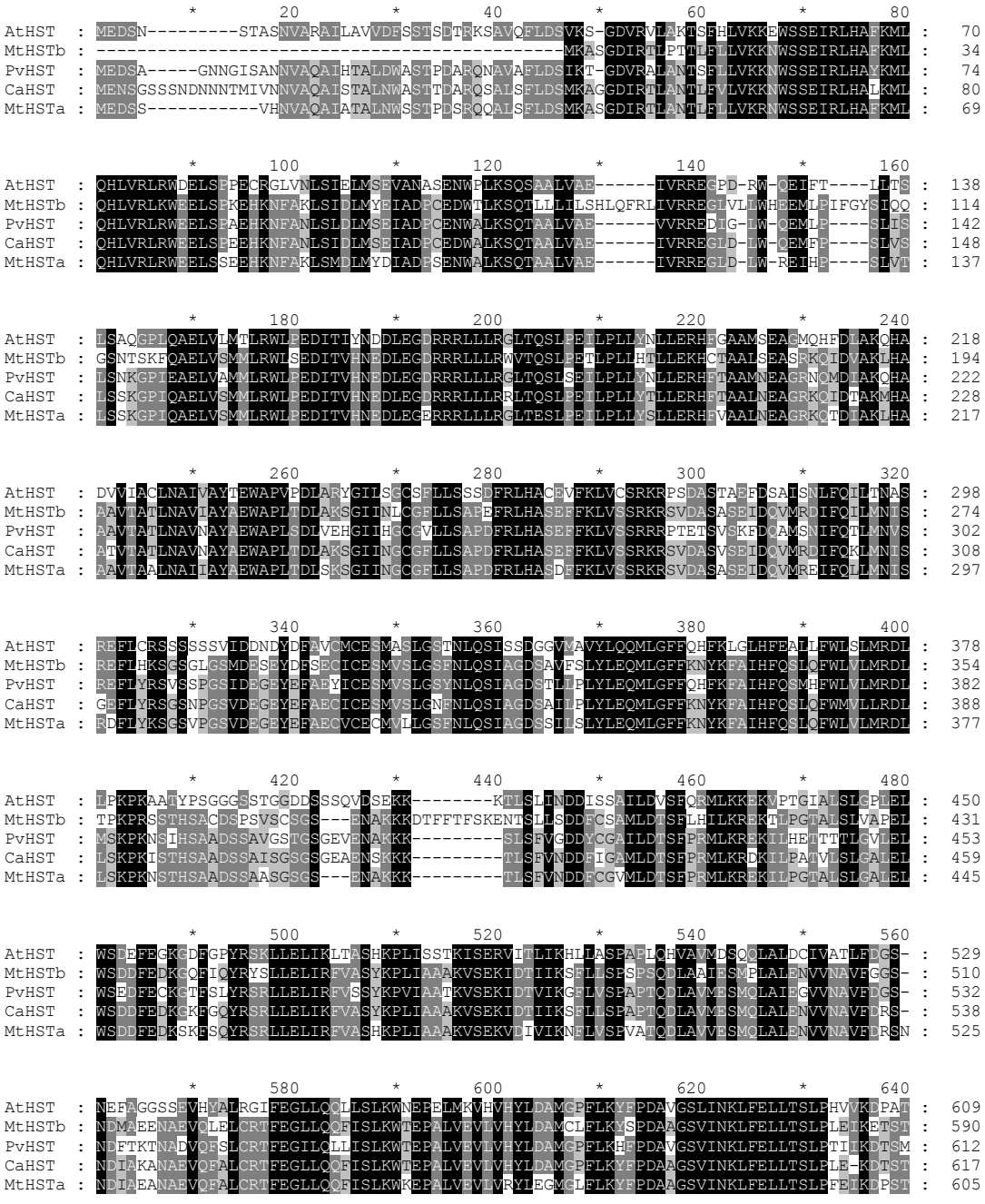
Appendix 3.13 Identity matrix (%) derived from the multiple sequence alignment of *SQUINT* (SQN) genes in the species *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca), *Medicago truncatula* (Mt) and *Phaseolus vulgaris* (Pv). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

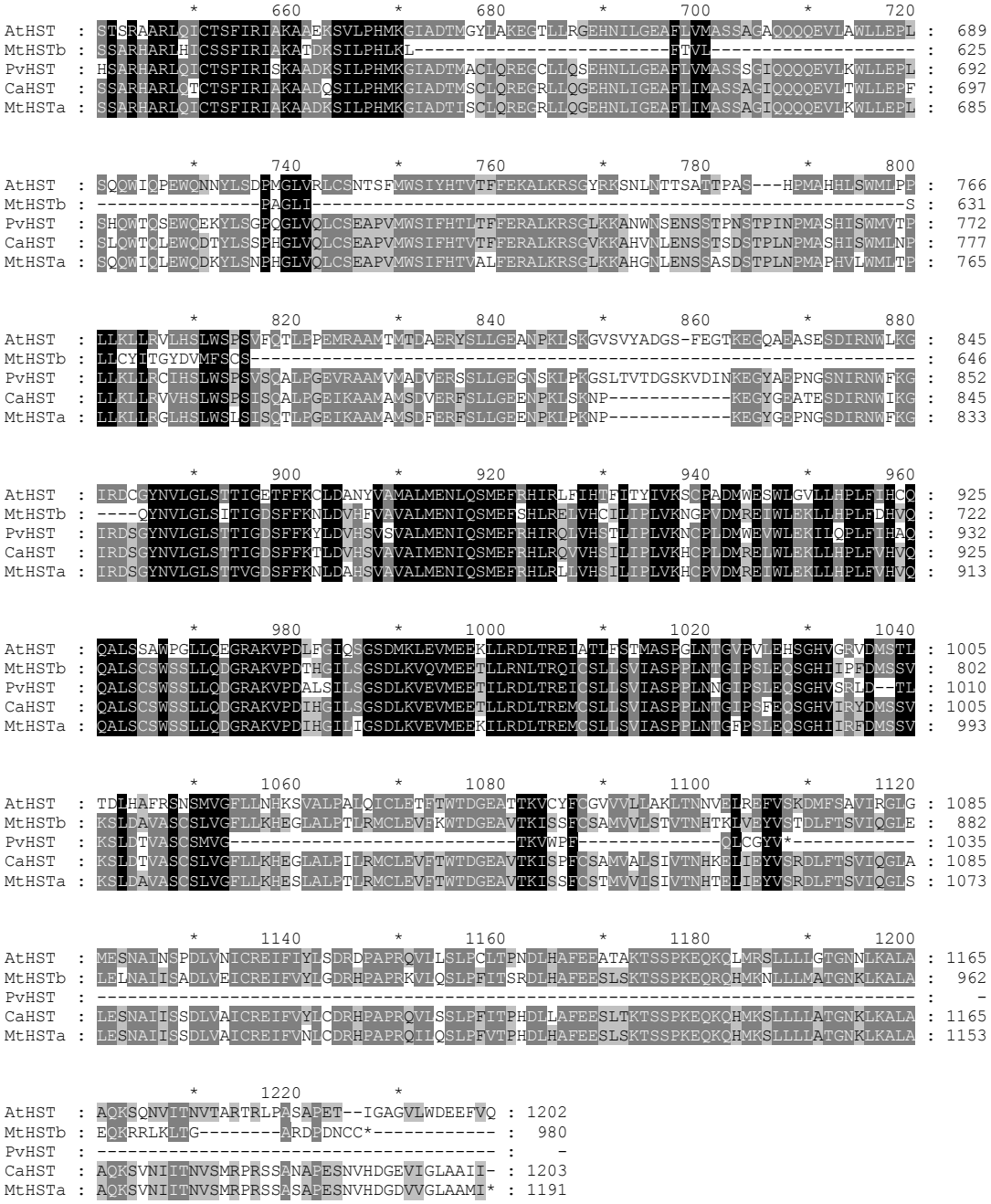
	Accession	AtSQN	CaSQNa	CaSQNb	MtSQNa	MtSQNb	PvSQNa	PvSQNb
AtSQN	AT2G15790		75.9	74.5	76.5	71.2	74.8	79.2
CaSQNa	LOC101491270	75.9		77.6	90.0	74.8	77.6	86.7
CaSQNb	LOC101510309	74.5	77.6		75.6	86.9	83.1	78.4
MtSQNa	Medtr4g086760	76.5	90.0	75.6		73.8	76.8	88.1
MtSQNb	Medtr8g079690	71.2	74.8	86.9	73.8		80.6	76.0
PvSQNa	Phvul.003G246400	74.8	77.6	83.1	76.8	80.6		77.9
PvSQNb	Phvul.003G294200	79.2	86.7	78.4	88.1	76.0	77.9	



Appendix 3.14 Identity matrix (%) derived from the multiple sequence alignment of *HASTY* (HST) genes in the species *Arabidopsis thaliana* (At), *Phaseolus vulgaris* (Pv), *Cicer arietinum* (Ca) and *Medicago truncatula* (Mt). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

	Accession	AtHST	MtHSTb	PvHST	CaHST	MtHSTa
AtHST	AT3G05040		46.9	60.1	62.5	61.7
MtHSTb	Medtr1g063040	46.9		54.3	65.8	65.4
PvHST	Phvul.006G097400	60.1	54.3		75.6	73.1
CaHST	LOC101489604	62.5	65.8	75.6		86.9
MtHSTa	Medtr7g114260	61.7	65.4	73.1	86.9	





Appendix 3.15 Maximum parsimony tree derived from the alignment of *DICER-like* (*DCL*) genes in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Phaseolus vulgaris* (Pv) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions in Table 1 were aligned using ClustalW and PAUP* software was used to build the tree. Numbers in branches represent bootstrap consensus support from 1000 replications.

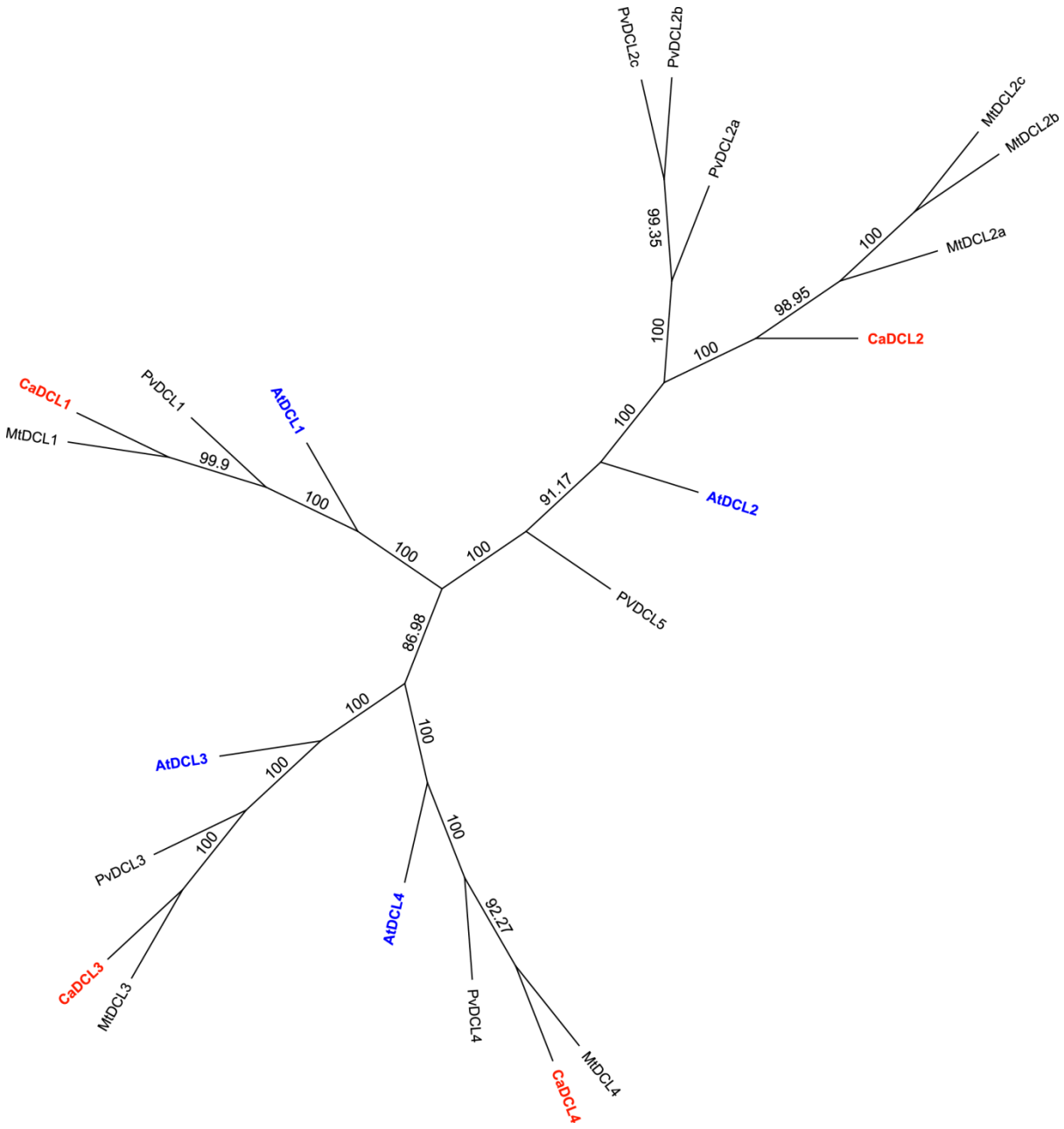


Table 1 Accesion number of the sequences used in the alignment and tree construction of *DCL* genes in four plant species.

Arabidopsis thaliana	
DCL1	AT1G01040
DCL2	AT3G03300
DCL3	AT3G43920
DCL4	AT5G20320

Cicer arietinum	
DCL1	LOC101502324
DCL2	LOC101501625
DCL3	LOC101496220
DCL4	LOC101502793

Medicago truncatula	
DCL1	Medtr7g118350
DCL2a	Medtr2g030490
DCL2b	Medtr1g060740
DCL2c	Medtr8g069975
DCL3	Medtr3g105390
DCL4	Medtr4g116860

Phaseolus vulgaris	
DCL1	Phvul.009G260000
DCL2a	Phvul.006G127100
DCL2b	Phvul.006G127200
DCL2c	Phvul.008G129500
DCL3	Phvul.009G083800
DCL4	Phvul.003G175700
DCL5	Phvul.006G111800

Appendix 3.16 Maximum parsimony tree derived from the alignment of *ARGONAUTE* (*AGO*) genes in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Phaseolous vulgaris* (Pv) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions in Table 1 were aligned using MUSCLE and PAUP* software was used to build the tree. Numbers in branches represent bootstrap consensus support from 1000 replications.

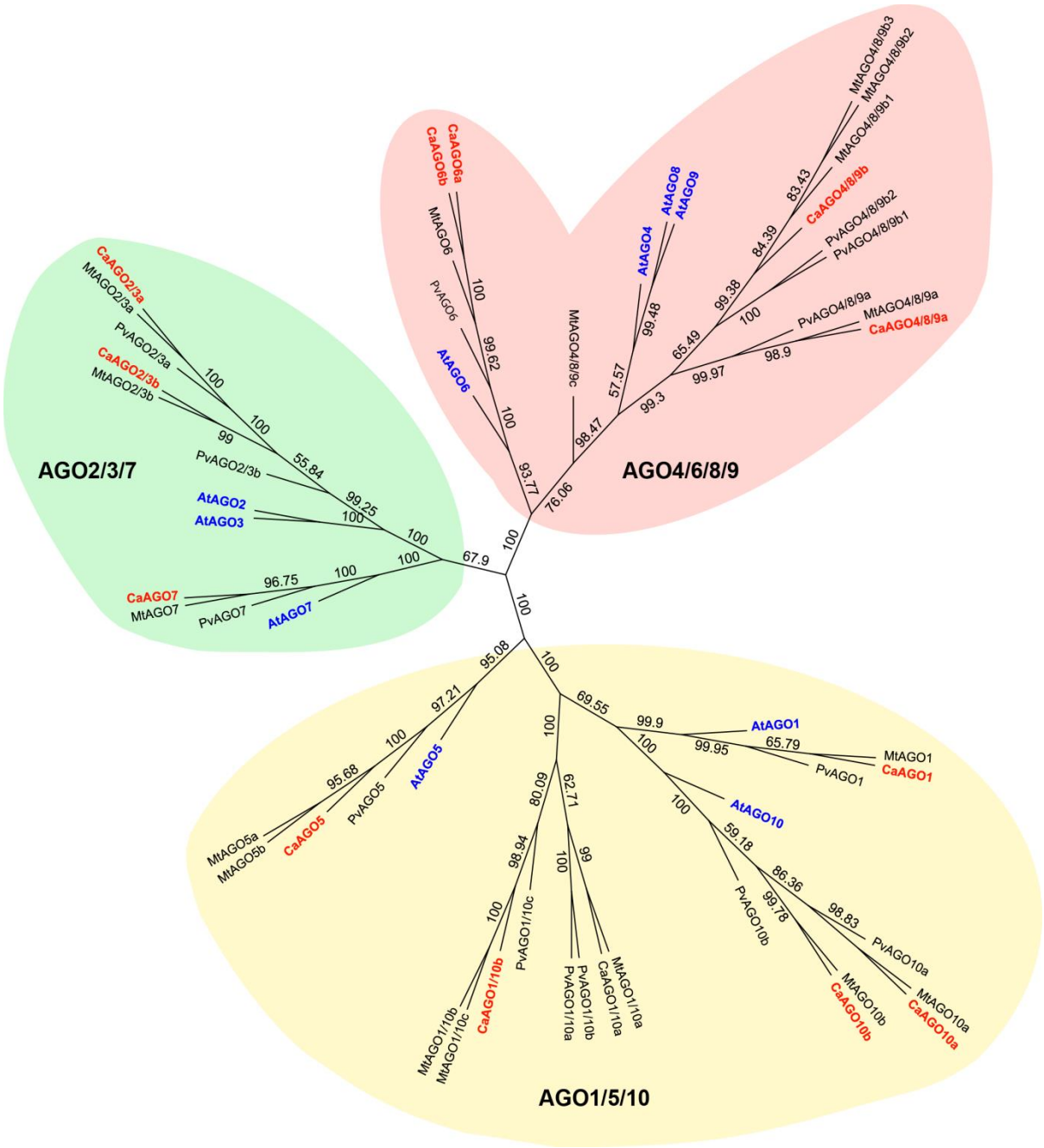


Table 1 Accession number of the genes used for the alignment and tree construction of AGO genes in four plant species.

Arabidopsis thaliana	
AGO1	AT1G48410
AGO2	AT1G31280
AGO3	AT1G31290
AGO4	AT2G27040
AGO5	AT2G27880
AGO6	AT2G32940
AGO7	AT1G69440
AGO8	AT5G21030
AGO9	AT5G21150
AGO10	AT5G43810

Medicago truncatula	
AGO1	Medtr6g477980
AGO1/10a	Medtr4g113200
AGO1/10b	Medtr2g059590
AGO1/10c	Medtr2g059790
AGO2/3a	Medtr2g028910
AGO2/3b	Medtr4g083610
AGO4/8/9a	Medtr3g078647
AGO4/8/9b1	Medtr5g087870
AGO4/8/9b2	Medtr5g087890
AGO4/8/9b3	Medtr6g043180
AGO4/8/9c	Medtr3g010650
AGO5a	Medtr4g056430
AGO5b	Medtr4g056470
AGO6	Medtr3g083300
AGO7	Medtr5g042590
AGO10a	Medtr1g102630
AGO10b	Medtr1g041345

Phaseolus vulgaris	
AGO1	Phvul.004G142900
AGO1/10a	Phvul.003G160200
AGO1/10b	Phvul.003G160000
AGO1/10c	Phvul.009G199500
AGO2/3a	Phvul.006G131700
AGO2/3b	Phvul.002G100100
AGO4/8/9a	Phvul.006G021200
AGO4/8/9b1	Phvul.008G206600
AGO4/8/9b2	Phvul.008G206500
AGO5	Phvul.011G088200
AGO6	Phvul.011G169400
AGO7	Phvul.003G046700
AGO10a	Phvul.007G062800
AGO10b	Phvul.007G278600

Cicer arietinum	
AGO1	LOC101511115
AGO1/10a	LOC101498544
AGO1/10b	LOC101509023
AGO2/3a	LOC101515466
AGO2/3b	LOC101508479
AGO4/8/9a	LOC101510486
AGO4/8/9b	LOC101514212
AGO5	LOC101513622
AGO6a	LOC101496222
AGO6b	LOC101502569
AGO7	LOC101488825
AGO10a	LOC101498267
AGO10b	LOC101499992

Appendix 3.17 Maximum parsimony tree derived from the alignment of *Giberellin oxidase* genes (*GAox*) in *Cicer arietinum* (Ca, in red), *Pisum sativum* (Ps) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions in table 1 were aligned using MUSCLE algorithm, and PAUP* software was used to build the tree. Numbers in branches represent bootstrap consensus support from 1000 repetitions.

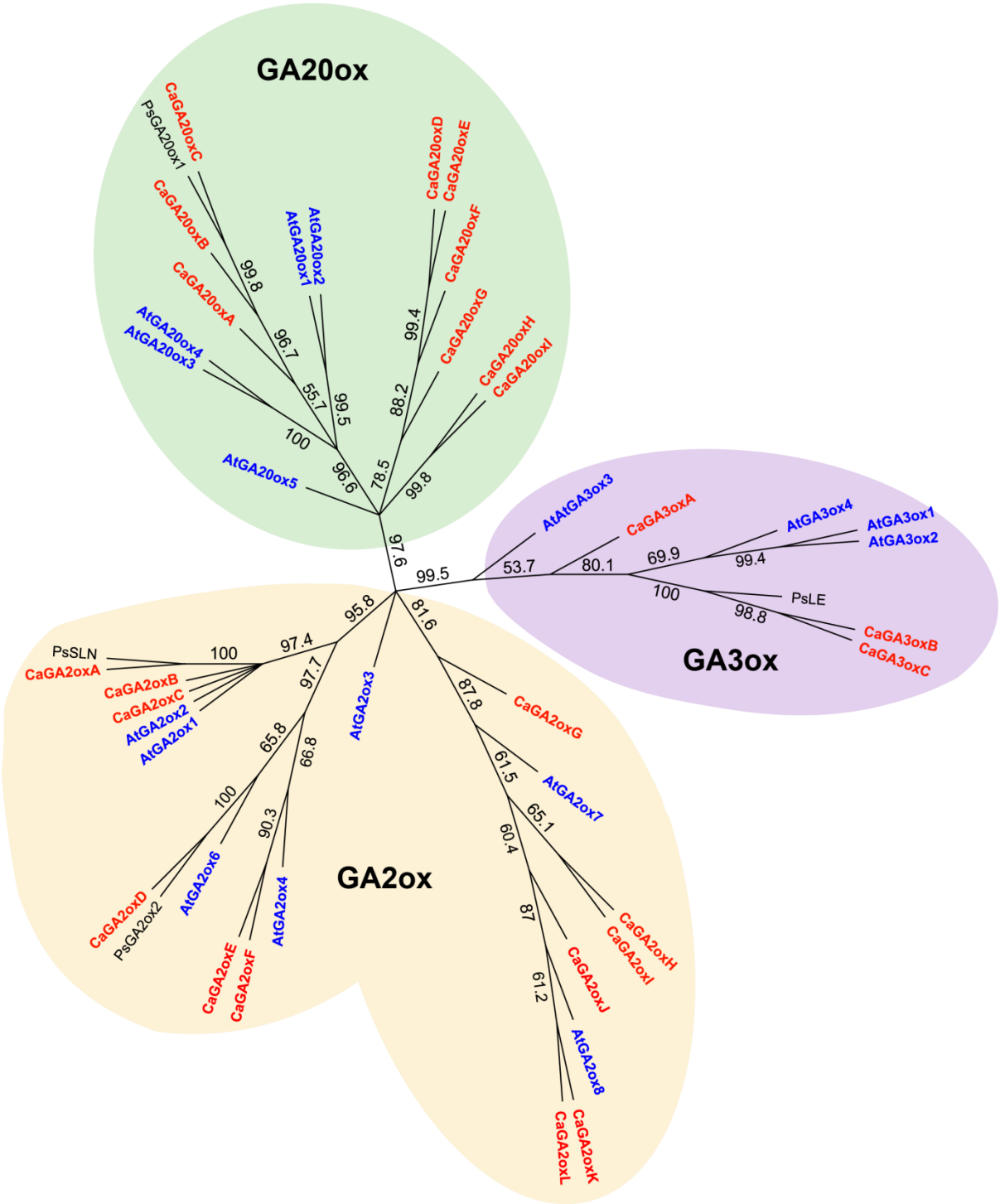


Table 1 Accession number of the protein sequences used in the alingment and tree construction of GAox genes in 4 plant species

Arabidopsis thaliana	
GA2ox1	AT1G78440
GA2ox2	AT1G30040
GA2ox3	AT2G34550
GA2ox4	AT1G47990
GA2ox6	AT1G02400
GA2ox7	AT1G50960
GA2ox8	AT4G21200
GA3ox1	AT1G15550
GA3ox2	AT1G80340
GA3ox3	AT4G21690
GA3ox4	AT1G80330
GA20ox1	AT4G25420
GA20ox2	AT5G51810
GA20ox3	AT5G07200
GA20ox4	AT1G60980
GA20ox5	AT1G44090

Pisum sativum	
SLN	AF101383
GA2ox2	AF100954
LE	U85045
GA20ox1	AF138704

Cicer arietinum	
GA2oxA	LOC101494727
GA2oxB	LOC101513626
GA2oxC	LOC101488778
GA2oxD	LOC101497057
GA2oxE	LOC101491391
GA2oxF	LOC101510182
GA2oxG	LOC101504070
GA2oxH	LOC101510845
GA2oxI	LOC101511169
GA2oxJ	LOC101514491
GA2oxK	LOC101515302
GA2oxL	LOC101502300
GA3oxA	LOC101503293
GA3oxB	LOC101498534
GA3oxC	LOC101498875
GA20oxA	LOC101491937
GA20oxB	LOC101490587
GA20oxC	LOC101491320
GA20oxD	LOC101500571
GA20oxE	LOC101492898
GA20oxF	LOC101492568
GA20oxG	LOC101492441
GA20oxH	LOC101488618
GA20oxI	LOC101491889

Appendix 3.18 Maximum parsimony tree derived from the alignment of vernalization-related genes in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Pisum sativum* (Ps) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions in Table 1 were aligned using ClustalW and PAUP* software was used to build the tree. Numbers in branches represent bootstrap consensus support from 1000 replications.

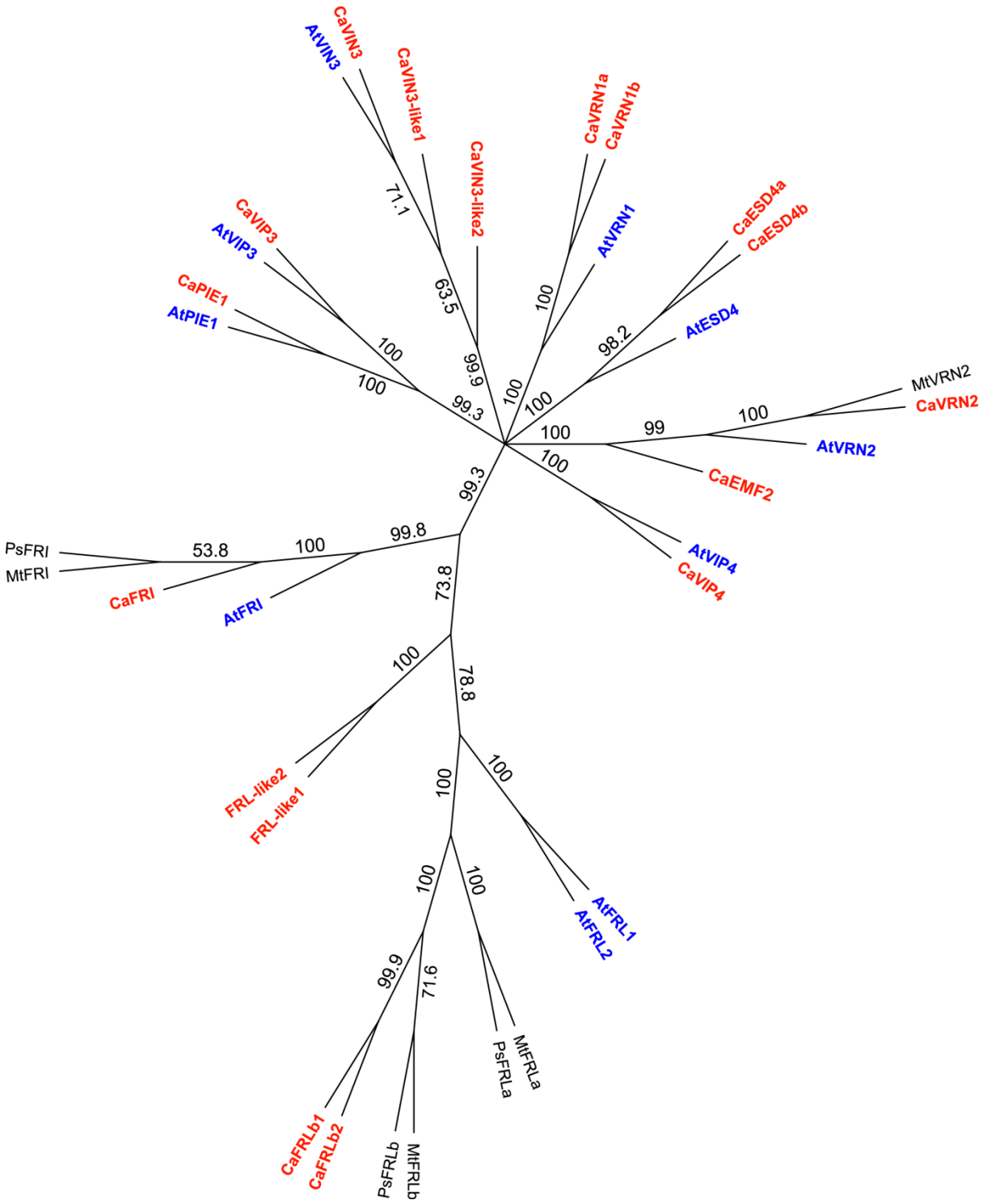


Table 1 Accession numbers of the sequences used in the alignment and tree construction of vernalization-related genes in four plant species.

Arabidopsis thaliana	
VRN1	AT3G18990
VIN3	AT5G57380
FRI	AT4G00650
FRL1	AT5G16320
FRL2	AT1G31814
VIP3	AT4G29830
VIP4	AT5G61150
ESD4	AT4G15880
PIE1	AT3G12810
VRN2	AT4g16845

Medicago truncatula	
FRI	Medtr3g098290
FRLa	Medtr3g056070
FRLb	Medtr5g094400
VRN2	Medtr5g013150

Pisum sativum	
FRI	PsCam036412
FRLa	PsCam035597
FRLb	PsCam000193

Cicer arietinum	
VRN1a	LOC101488634
VRN1b	LOC101507442
VIN3	LOC101498781
VIN3-like1	LOC101491891
VIN3-like2	LOC101505288
FRI	LOC101515253
FRLb1	LOC105852246
FRLb2	LOC101501660
FRL-like1	LOC101499213
FRL-like2	LOC101499533
VIP3	LOC101502792
VIP4	LOC101508547
ESD4a	LOC101514190
ESD4b	LOC101494627
PIE1	LOC101490716
VRN2	LOC101513120
EMF2	LOC101511538

Appendix 3.19 Accession number of CONSTANS and CONSTANS-like (COL) sequences used in the alignment and tree construction in five plant species.

Medicago truncatula	
COLa	Medtr7g018170
COLb	Medtr1g013450
COLc	Medtr3g105710
COLd	Medtr4g128930
COLe	Medtr3g082630
COLf	Medtr5g069480
COLg	Medtr7g108150
COLh	Medtr7g083540
COLj	Medtr2g088900
COLi	Medtr8g104190
COLk	Medtr1g110870

Cicer arietinum	
COLa	LOC101489582
COLb	LOC101512206
COLc	LOC101490500
COLd	LOC101510742
COLe	LOC101498290
COLf	LOC101491305
COLg	LOC101499146
COLh	LOC101504031
COLj	LOC101499105
COLi	LOC101506376
COLk	LOC101499877

Pisum sativum	
COLa	PsCam010875
COLb	PsCam000878
COLc	PsCam057979
COLd	PsCam001031
COLe	PsCam045654
COLf	PsCam042897
COLg	PsCam045953
COLh	PsCam010810
COLj	PsCam036791
COLi	PsCam042996
COLk	PsCam020962

Glycine max	
COL1a	Glyma.08G255200
COL2b	Glyma.19G039000
COL3b	Glyma.06G059600
COL4a	Glyma.13G093800
COL8a	Glyma.02G223700
COL9a	Glyma.13G009300
COL11a	Glyma.03G209800
COL12a	Glyma.02G152900
COL13a	Glyma.16G050900
COL10a	Glyma.12G196100
COL5a	Glyma.07G091400
COL6a	Glyma.05G233700
COL7a	Glyma.10G274300
COL1b	Glyma.18G278100
COL2a	Glyma.13G050300
COL3a	Glyma.04G058900
COL4b	Glyma.17G066600
COL8b	Glyma.14G190400
COL9b	Glyma.20G060400
COL11b	Glyma.19G207100
COL12b	Glyma.10G021400
COL13b	Glyma.19G099700
COL10b	Glyma.13G306400
COL5b	Glyma.09G184600
COL6b	Glyma.08G041100
COL7b	Glyma.20G115600

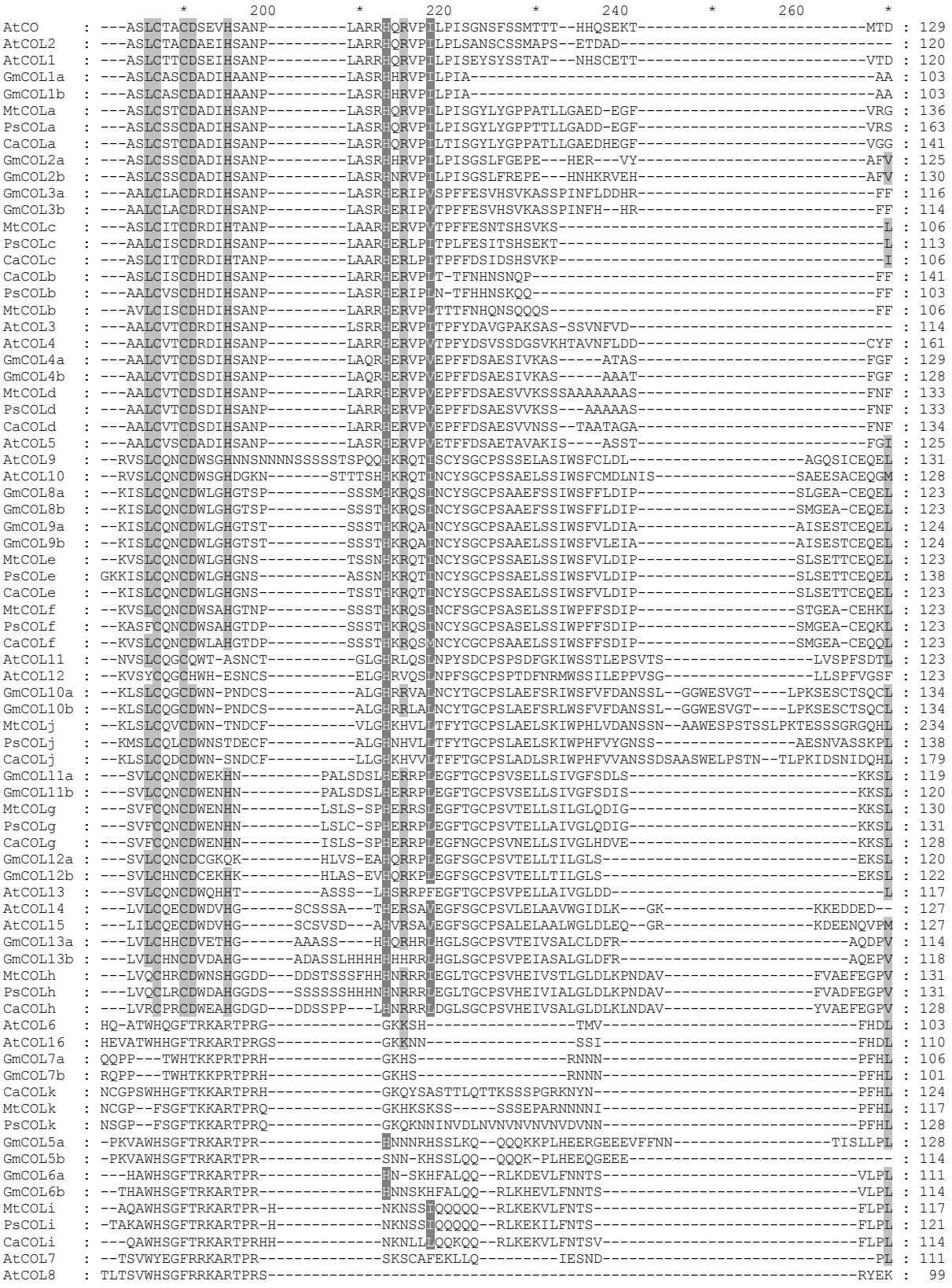
Arabidopsis thaliana	
CO	AT5G15840
COL1	AT5G15850
COL2	AT3G02380
COL3	AT2G24790
COL4	AT5G24930
COL5	AT5G57660
COL6	AT1G68520
COL7	AT1G73870
COL8	AT1g49130
COL9	AT3G07650
COL10	AT5G48250
COL11	AT4G15250
COL12	AT3G21880
COL13	AT2G47890
COL14	AT2G33500
COL15	AT1G28050
COL16	AT1G25440

Sequence alignment for Figure 3.1. Protein sequences of accessions listed in the table on previous page were aligned with MUSCLE in Geneious 8.0.

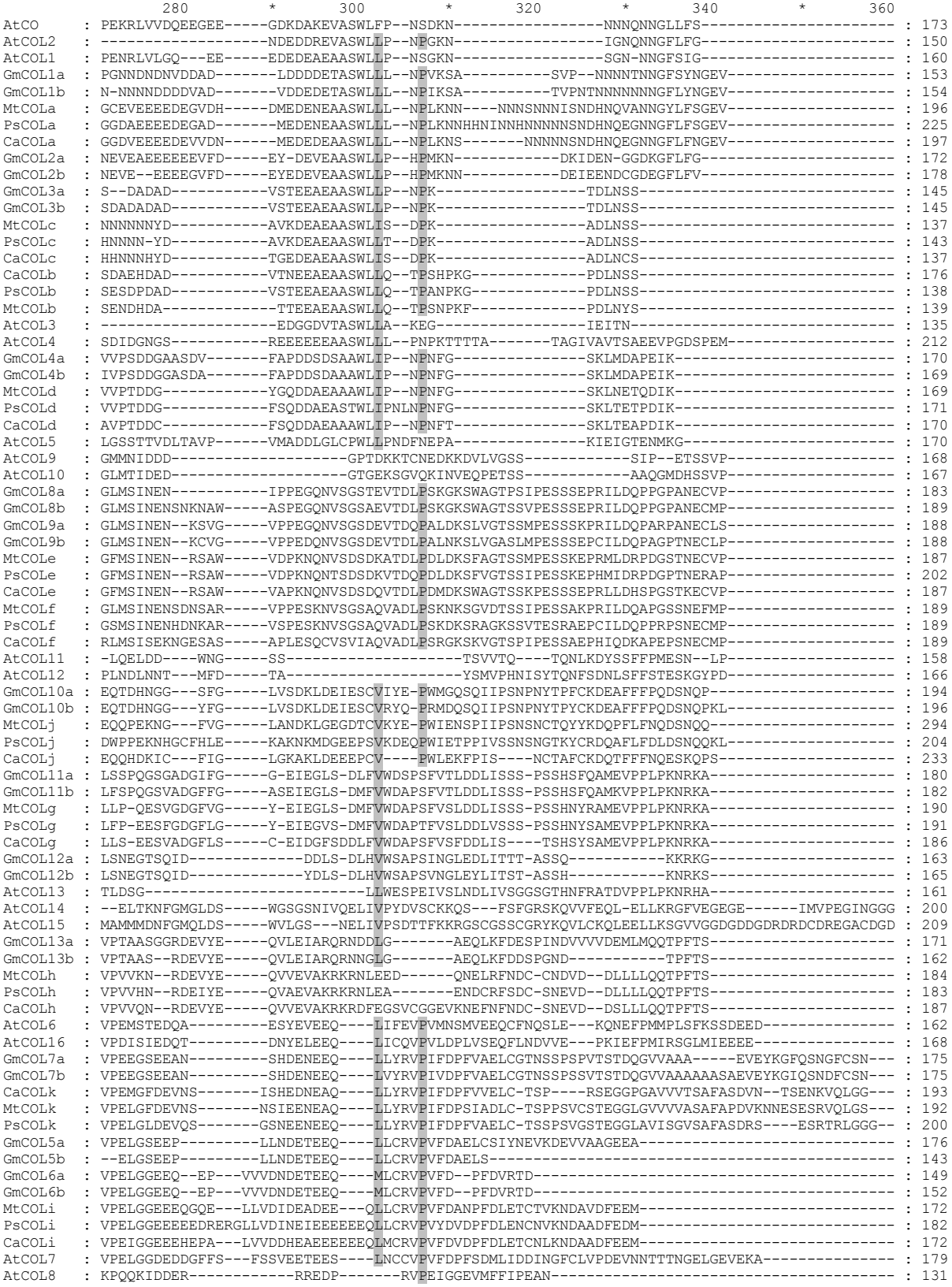
```
                  *        20         *        40         *        60         *        80         *
AtCO     : ------------------------------------------------------------------------MLKQESNDIGSG------- :  12
AtCOL2   : ------------------------------------------------------------------------MLKEESNESGT-------- :  11
AtCOL1   : ------------------------------------------------------------------------MLKVESN------------ :   7
GmCOL1a  : ------------------------------------------------------------------------MLDG--------------- :   4
GmCOL1b  : ------------------------------------------------------------------------MLEG--------------- :   4
MtCOLa   : ------------------------------------------------------------------------MLEQDFLTTTSA----TAT :  15
PsCOLa   : ----------------------------------------MKLLHFNILNTVCPKLTNKPFITHSLTPQMLEQKFLTTTS------AT :  42
CaCOLa   : ------------------------------------------------------------------------MLEQELFTTTTTSATATAT :  19
GmCOL2a  : ------------------------------------------------------------------------MLKEGTNN--------VGG :  11
GmCOL2b  : ------------------------------------------------------------------------MLKEGTNN--------VGG :  11
GmCOL3a  : ------------------------------------------------------------------------------------------- :   -
GmCOL3b  : ------------------------------------------------------------------------------------------- :   -
MtCOLc   : ------------------------------------------------------------------------------------------- :   -
PsCOLc   : ------------------------------------------------------------------------------------------- :   -
CaCOLc   : ------------------------------------------------------------------------------------------- :   -
CaCOLb   : ------------------------------------------------------MPYPHNHNHTSIVTLPIQFNTLTTLSTFKATLNS---- :  34
PsCOLb   : ------------------------------------------------------------------------------------------- :   -
MtCOLb   : ------------------------------------------------------------------------------------------- :   -
AtCOL3   : ------------------------------------------------------------------------------------------- :   -
AtCOL4   : ------------------------------------------------MDPTWIDSLTRSCEANSNTNHKRKRERETLKHREKKKKRF---- :  40
GmCOL4a  : ------------------------------------------------------------------------MGIER---GGLKGFR---- :  12
GmCOL4b  : ------------------------------------------------------------------------MGIER---GGFKGFR---- :  12
MtCOLd   : ------------------------------------------------------------------------MGIER---GGLKSLR---- :  12
PsCOLd   : ------------------------------------------------------------------------MGIERAGGGGLKSFR---- :  15
CaCOLd   : ------------------------------------------------------------------------MGIERGGGGGLKGFR---- :  15
AtCOL5   : ------------------------------------------------------------------------MGFGL---ESIKSIS---- :  12
AtCOL9   : ------------------------------------------------------------------------------------------- :   -
AtCOL10  : ------------------------------------------------------------------------------------------- :   -
GmCOL8a  : ------------------------------------------------------------------------------------------- :   -
GmCOL8b  : ------------------------------------------------------------------------------------------- :   -
GmCOL9a  : ------------------------------------------------------------------------------------------- :   -
GmCOL9b  : ------------------------------------------------------------------------------------------- :   -
MtCOLe   : ------------------------------------------------------------------------------------------- :   -
PsCOLe   : ------------------------------------------------------------------------------------------- :   -
CaCOLe   : ------------------------------------------------------------------------------------------- :   -
MtCOLf   : ------------------------------------------------------------------------------------------- :   -
PsCOLf   : ------------------------------------------------------------------------------------------- :   -
CaCOLf   : ------------------------------------------------------------------------------------------- :   -
AtCOL11  : ------------------------------------------------------------------------------------------- :   -
AtCOL12  : ------------------------------------------------------------------------------------------- :   -
GmCOL10a : ------------------------------------------------------------------------------------------- :   -
GmCOL10b : ------------------------------------------------------------------------------------------- :   -
MtCOLj   : MGLSDDRGEYIGAKTTSFKGLPQSNVEEAKGLKEAINCLGNLRFPSLSIKLDWGFHILGLDKEQSSSDYRSCVPSNPIHSDPIHSISFRF :  90
PsCOLj   : ------------------------------------------------------------------------------------MVISVIF :   7
CaCOLj   : ------------------------------------------------------MSHSTFNLSCYYFMHPPPFQSSNNNFFEGGGIQLFLVW :  38
GmCOL11a : ------------------------------------------------------------------------------MSG-------- :   3
GmCOL11b : ------------------------------------------------------------------------------MSGA------- :   4
MtCOLg   : ------------------------------------------------------------------------------MGGSPRNPNP :  10
PsCOLg   : ------------------------------------------------------------------------------MDGSPANPNP :  10
CaCOLg   : ------------------------------------------------------------------------------MGGSNKNEQE :  10
GmCOL12a : ------------------------------------------------------------------------------MMNGSP----- :   6
GmCOL12b : ------------------------------------------------------------------------------MMSGSPS---- :   7
AtCOL13  : ------------------------------------------------------------------------------MEAEEG---- :   6
AtCOL14  : ------------------------------------------------------------------------------------------M :   1
AtCOL15  : ------------------------------------------------------------------------------------------- :   -
GmCOL13a : ------------------------------------------------------------------------------------------- :   -
GmCOL13b : ------------------------------------------------------------------------------------------- :   -
MtCOLh   : ------------------------------------------------------------------------------------------- :   -
PsCOLh   : ------------------------------------------------------------------------------------------- :   -
CaCOLh   : ------------------------------------------------------------------------------------------- :   -
AtCOL6   : ------------------------------------------------------------------------------MMK---SLAS :   7
AtCOL16  : ------------------------------------------------------------------------------MMK---SLAN :   7
GmCOL7a  : ------------------------------------------------------------------------MSSATK---NAAN :  10
GmCOL7b  : ------------------------------------------------------------------------MSSATK---NAAN :  10
CaCOLk   : ------------------------------------------------------------------------------MTK---NVAN :   7
MtCOLk   : ------------------------------------------------------------------------MTCSSK---NVAN :  10
PsCOLk   : ------------------------------------------------------------------------MTCSSE---NLAK :  10
GmCOL5a  : ------------------------------------------------------------------------MTNEMK----EAS :   9
GmCOL5b  : ------------------------------------------------------------------------MTNEMK----EAS :   9
GmCOL6a  : ------------------------------------------------------------------------MRDMK----DAG :   8
GmCOL6b  : ------------------------------------------------------------------------MRDMK----DAG :   8
MtCOLi   : ------------------------------------------------------------------------MIIDMK-GDADAG :  12
PsCOLi   : ------------------------------------------------------------------------MITDMKSGGADAG :  13
CaCOLi   : ------------------------------------------------------------------------------MK-GDADSG :   8
AtCOL7   : ------------------------------------------------------------------------MVVDVES--RTAS :  11
AtCOL8   : ------------------------------------------------------------------------------MFCAEIMISK :  10
```

Sequence alignment for Figure 3.1. Continued

Sequence alignment for Figure 3.1. Continued



Sequence alignment for Figure 3.1. Continued



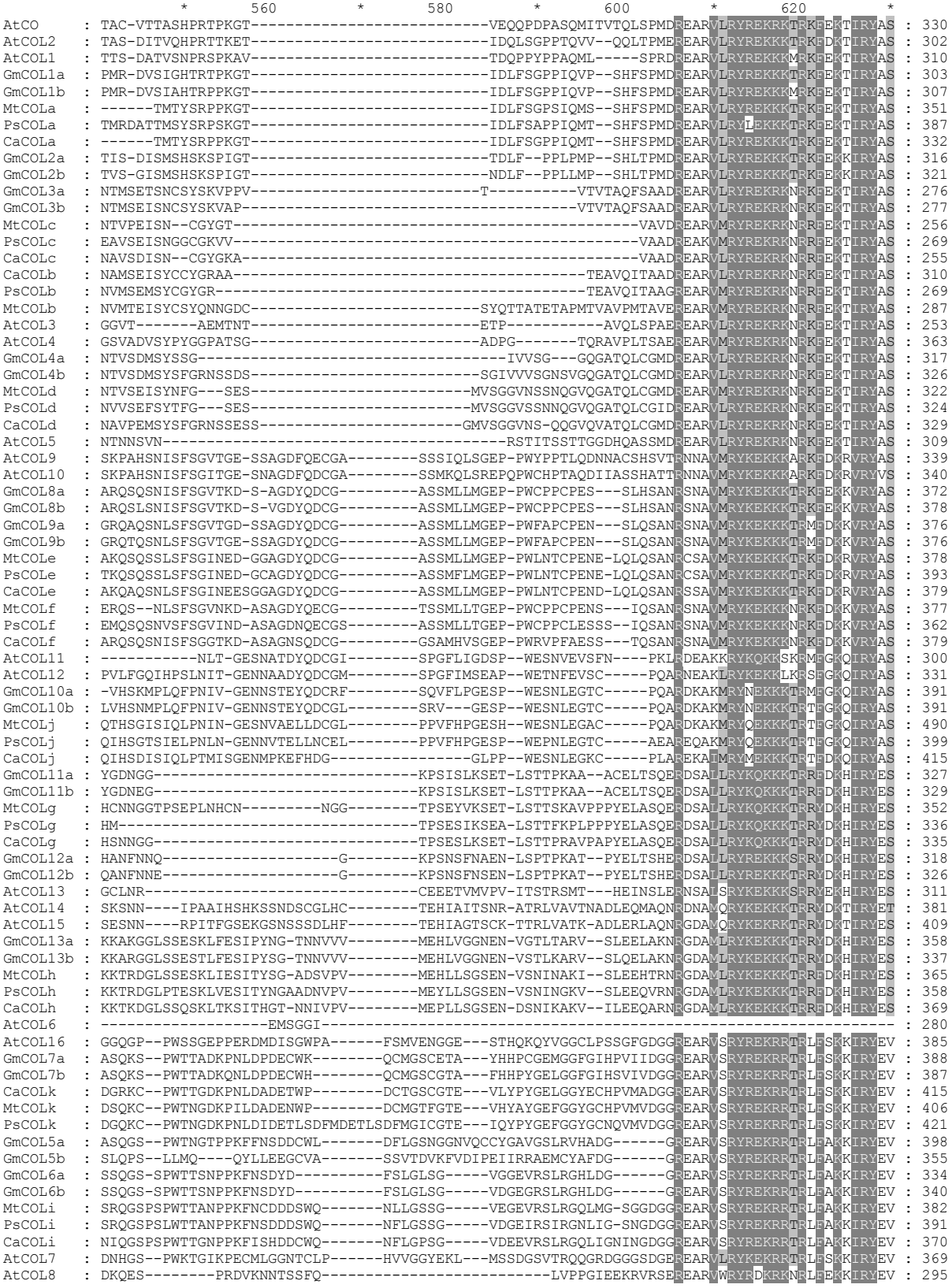
Sequence alignment for Figure 3.1. Continued

```
                  *      380        *       400        *       420        *       440        *
AtCO     : ----------DEYLN-LVDYNSSM---DYKFTG-EYSQHQQNC--------------------------------SVPQTSYGGDRVV : 214
AtCOL2   : ----------VEYLD-LVDYSSSM---DNQFEDNQYTHYQR-----------------------------------SFGGDGVV : 185
AtCOL1   : ----------DEFLN-LVDYSSS----DKQFTD-QSNQYQLDC--------------------------------NVPQRSYGEDGVV : 200
GmCOL1a  : ----------DEYLD-LVDDC----D-NHHFASVATTTDHYS----HQHQ-----------------------HFGVVSHKSYAGDSVV : 199
GmCOL1b  : ----------DEYLD-LVDNCNSCGD-NNHFASAAATTDHY-----AQHQ-----------------------HFAGVSQKSYAGDSVV : 203
MtCOLa   : ----------DEYLD-LVDCNSWGGD-ENTFTTNNTHHHDEYS-QQQQQQ-----------------------DHYGVPQKSYVGDSVV : 249
PsCOLa   : ----------DEYLD-LVDCNSCGGG-ENTFTTNNTHDHDYSRDQQQQRQ-----------------------DHYGVPQKNYVGDSVV : 279
CaCOLa   : ----------DEYLD-LVDCNSCGGGGDNNFNTNTTHHDEYS-------------------------------------------VV : 230
GmCOL2a  : ----------DEYFDNLVD-CNSCGHNNNQFSNVYDQHQQNYS--------------------------------NTVPQN----YAVV : 214
GmCOL2b  : ----------DEYLDNLVDCCNSCGHNDNQFSNVY-QHQQNY---------------------------------NTVPQN----YVVV : 219
GmCOL3a  : --------------QYLFSETE--PVPYIDLDYAAVDPKAEQ-----------------------------------KSSATADGVV : 181
GmCOL3b  : --------------QYLFSETE--PVPYIDLDYAAMDPKTEQ-----------------------------------KSSATADGVV : 181
MtCOLc   : --------------PYLFSDSE--AIPFMDLDYGVIEHKN-------------------------------------------DGVV : 165
PsCOLc   : --------------PYMFSDSE--AIPFMDLDYGVIEHKNV------------------------------------------DGVV : 172
CaCOLc   : --------------PYIFSDSE--AIPFMDLDYGTIDNKS--------------------------------------------DVVV : 165
CaCOLb   : --------------HYSFHEIDNAGIDFVSVDAKTDSPEQH-----------------------------------SPGGTADGVV : 213
PsCOLb   : --------------HYSFTEIDATDLNFVCVDAKTDSPEQH-----------------------------------SPG-TADGVV : 174
MtCOLb   : --------------HYSYPEID----DFVTVNTKTDLPEQN-----------------------------------SPGTTADGVV : 172
AtCOL3   : ----------------LFS----------DLDYPKIEVTSE-------------------------------------ENSSGNDGVV : 160
AtCOL4   : ----------NTGQQYLFSDPD----PYLDLDYGNVDPKVESL-----------------------------------EQNSSGTDGVV : 252
GmCOL4a  : ------------SKEIFFSEMD----PFLDFDYSNSFQNH-------------------------------------NSAVNDSVV : 203
GmCOL4b  : ------------SKEIFFSEMD----PFLDFDYSNSFQNN-------------------------------------NSAGNDSVV : 202
MtCOLd   : ------------TREMFFSDMD----PFLDFDYSNNFQNN---------------------------------------NCSNAMNDSVV : 204
PsCOLd   : ------------TRDMFFSDMD----PFLDFDYSNNFQNN---------------------------------------NCN-AINDSVV : 205
CaCOLd   : ------------TRDMFFSDMD----PFIDFDYSNTFQNNN--------------------------------------NCSSAINDSVV : 206
AtCOL5   : ------------SSDFMFSDFD----RLIDFEFPNSFNHH---------------------------------------QNNAGGDSLV : 204
AtCOL9   : --------QGKSSSAKDVGMCED--DFYGNLGMDEVDMALENYEE---------------------------------LFGTAFNPSEELF : 216
AtCOL10  : --------E-NSSMAKELGVCED--DFNGNLISDEVDLALENYEE---------------------------------LFGSAFNSSRYLF : 214
GmCOL8a  : --------KLYCPGKKASGICED-DDLYDDFIMDEVDLELENYEE---------------------------------LFGMALSHSEELF : 232
GmCOL8b  : --------KLYCPGKKVSGIYED-DDLYDDFIMDEVDLQLENYEE---------------------------------LFGMALSHSEELF : 238
GmCOL9a  : --------KLYCPATKFPALCEN-DNLYDDFNMDEVDLNLENYEE---------------------------------LFGMALSHSEELF : 237
GmCOL9b  : --------KLYCPATKCPALSED-DNLYDDFNMDEVDLDLENYEE---------------------------------LFGMALSHSEELF : 237
MtCOLe   : --------KLYCPATNCREASDDDDDLYGDFDMDEMDINMENYDE---------------------------------LFGMALTHSEELF : 237
PsCOLe   : --------KLYCPATKCPEISEDDDDLYGDFDMDEMDINMENYDE---------------------------------LFGMALTHSEELF : 252
CaCOLe   : --------KLYCSATKCPEEREDKDDLYGDFDMDEMDINIENYDE---------------------------------LFGMALTHSEELF : 237
MtCOLf   : --------KLLCPSRKAPALCED-DKLLDDFNIDEMDFELENYSE---------------------------------LFDFALNHSEEFF : 238
PsCOLf   : --------KLQCPGSTASALCED-DNLYDGFNIDDMDLELENYKE---------------------------------VFSYALNNSEEFF : 238
CaCOLf   : --------KLLCPSKKSPALCED-DNLYDDFNIDEMDLELENYKE---------------------------------LFGVDLSHTDEFL : 238
AtCOL11  : --------KVIEEE-------CSGLDLCEGINLDDAPLNFNASND---------------------------------IIGCS-SLDNTK- : 199
AtCOL12  : --------MVLKLE-------EGEEDLCEGLNLDDAPLNFDVGDD---------------------------------IIGCS-SEVHIEP : 208
GmCOL10a : --------KECPN-----LGIHDGSELCEGFNVDNLQLNFESADE---------------------------------IFGCSQDATRYHL : 239
GmCOL10b : --------QECPN-----LGIHDGNELCEGFNVDNVQLNFEIADE---------------------------------IFYCSQVATRYHL : 241
MtCOLj   : ----------------KDLIIHEGTSLCEGFNVDDIQLNFESADEI--------------------------------IFDCSQTATKYNH : 337
PsCOLj   : --------QGCPDV--KELVFHEGTSLCEGFNVDDVQLDFESADE---------------------------------IFGRSQSAAKYNH : 252
CaCOLj   : --------KICPDIN-KDEAIHEGTSLYEGLNVDDIQLNFETVNE---------------------------------ILDCSQTSTKYNQ : 282
GmCOL11a : --------ACGQHKEEILSQLRELAKSEP------LDLEPYV-----------------------------------SSGNLSSGF : 217
GmCOL11b : --------ACGQHKEEILSQLRELAKSEP------LDLEPYV-----------------------------------QSGNLSLGF : 219
MtCOLg   : --------ACGRHREEILNQLREMTKSEP------YDPEEYI-----------------------------------PPANLSTSF : 227
PsCOLg   : --------ACGRQKEEILSQLREMTKSEP------LDQEDYM-----------------------------------PSRSLSIDF : 228
CaCOLg   : --------ACGRHQEEILNQLREMVKSEP------LDPEEYM-----------------------------------QPKNLSTGF : 223
GmCOL12a : --------ACGRHKEEILSQLRELIKLEPDLIH----AERQG-----------------------------------QFGNLPTDF : 202
GmCOL12b : --------AFGRHKEEILSQLRELIKLEPDLIHGEVDAERQG-----------------------------------QFGNLPTGF : 208
AtCOL13  : --------TCGKYKDEMIRQLRGLSRSEP------GCLKFET-----------------------------------PDAEIDAGF : 198
AtCOL14  : S-------ISQPSPTTSFTSLLMSQSLCGNG-----MQWNAT--NH-------------------------------STGQNTQIWDFNL : 245
AtCOL15  : GDGEAGEGLMVPEMSERLKWSRDVEEINGGGGGGVNQQWNATTTNP-------------------------------SGGQSSQIWDFNL : 268
GmCOL13a : --------LLMLPNSESEFDSRKSNDNNGYGTEAGDLLWNYN-----------------------------------PAYQPPQVWDFQL : 218
GmCOL13b : --------LLMLQNSESEFDARKSNDNNGYGTEAWDLHWNYNN---------------------------------PAYQPPQVWDFQL : 210
MtCOLh   : --------LLNFS-SEFDVGVKR-NSN-DYGNESGLLLWDRN-------------------------------PSYQPPQVWDFQL : 228
PsCOLh   : --------LLNFS-SEFDVGAKK-INN-DYGNESGVLLWNRN-------------------------------PTYQPPQVWDFQL : 227
CaCOLh   : --------MLNWP-SEFEVEAKK-NNNNDYGTEPGVLLWDRN-------------------------------PTYQPPQVWDFQL : 232
AtCOL6   : ---------DDNAESCLNGLFP-TDMELAQFTADVETLLG-GGDRE----------------FHSIEELGLG--EMLKIEK----EEVE : 218
AtCOL16  : ----------DNAESCLNGFFP-TDMELEEFAADVETLLGRGLDTE----------------SYAMEELGLSNSEMFKIEKDEIEEEVE : 230
GmCOL7a  : ----------SSEMENLHGMLP-SDAELAEFAADVESLLGRGLENE----------------CVGMEELGLVDAKEEECS--VGSGKVK : 235
GmCOL7b  : ----------SNEIENLHGMLP-SDAELAEFAADVESFLGRGLENK----------------CVGMEELGLVDTKEEECS--VGSGKVK : 235
CaCOLk   : --------DSNYEMENFHGILP-SDIELAEFAADVESLLGRGLENE----------------CIGMEELGLVDTKREVYS----EGKVK : 253
MtCOLk   : --------DNNYEMESFHGLLP-SDIELAEFAADVESLLGRGLENE----------------CIGMEELGLIDTKHEESEKWKCGGKVK : 256
PsCOLk   : --------DSNYEMEGFHGLLP-SDIELAEFAADVESLLGRGLENE----------------CMGMEELGLIDAKYEDEE-WECSGKVK : 263
GmCOL5a  : -----------LDLENFSSEFLPSDTDLAEFTADVKSFLGNGADEDSPDDDEHVKESELLILDCKEEEGDEEMDECIDGVLIGA-KDAMV : 254
GmCOL5b  : -----------FDLENFSSEFLPTDMDLAEFSADVESLLGNGVDEDSPVK----KGSELVLLDCKEE-GDDEMDACVNG--IGANKDAMV : 215
GmCOL6a  : ------------DLDSFS------DMDFAEFAADVESLLDK-------------------------EDDEISACVGGG---EGVQGAMV : 192
GmCOL6b  : ------------DLDSFS------DMDFSEFAADVEGRLDK-------------------------EDDEISAYVGGG---EGVQGVLA : 195
MtCOLi   : -----------CDLDSFCEF----DVDLAEFAANVESLLGVGSSE----------------IQENSSGQVFDYKQEN-EMDASKSEML : 228
PsCOLi   : -----------CDLDSFCEF----DVDLAEFAANVESLLGVGSSE----------------IHENSTEQVFDYKQEN-EMDPSKNEIL : 238
CaCOLi   : -----------CDLDSFVEF----DVDLAEFAANVESLLGV----------------HENSTTQVFDYKQENNEMDASKNEML : 224
AtCOL7   : ----------IMDDEGFMGFVP-LDMDLEDLTMDVESLLE-------------------------EEQLCLGFKEPN-----DVGVIK : 226
AtCOL8   : ----------------------DDDMTSLVPEFEGFTEMGFFLSN--------------HNGTEETTKQFNFEEEADTMEDLYYNGE : 182
```

Sequence alignment for Figure 3.1. Continued

```
                460         *       480         *       500         *       520         *       540
AtCO     : PLKLEESRGH-----------------QCHNQQNFQFNI-KYG-SSGTHYNDN---------------GSINHN-AYISSMETGVVPES : 268
AtCOL2   : PLQVEESTSH-----------------LQQSQQNFQLGI-NYGFSSGAHYNNNS--------------LKDLNHS-ASVSSMDISVVPES : 242
AtCOL1   : PLQIEVSKG------------------MYQEQQNFQLSI-NCG-SWGALRSSN---------------GSLSHM-VNVSSMDLGVVPES : 253
GmCOL1a  : PVQHH-------------------------QHFQLGL-EFDNSKAAFSYNA----------------SVNQS-VSVSSMDIGVVPES : 243
GmCOL1b  : PVQQH-------------------------QHFQLGL-EFDNSKPAFSYNG----------------SVSQS-VSVSSMDIGVVPES : 247
MtCOLa   : PVQQQQ-----------------------VQNFQLGL-EFESSKAGFSYNGG--------------SISQS-VSVSSMDVGVVPES : 296
PsCOLa   : PVQQQH-----------------------LQNFQLGL-EFESSKAGFSYNGG--------------SISQS-VSVSSMDVGVVPES : 326
CaCOLa   : PVQQQQ-----------------------VQNFQLGLDQFESSKPGFSYNG----------------SISQS-VSVSSMDVGVVPES : 277
GmCOL2a  : PVQVP------------------------QHFQPGL-DFDSSKAGFSYDG----------------SLSQS-VSVSSMDVGVVLES : 258
GmCOL2b  : PVQVP------------------------QHFQPGL-DFDSSKAGFSYDG----------------SLSQS-VSVSSMDVGVVPES : 263
GmCOL3a  : PVQSNFE----------------------PFAYGYKYN--------TTLS--QS--------------Q--MSQSVSS-SSMEVGVVPDG : 222
GmCOL3b  : PVQSNFE----------------------PFTYGYKYN--------TTLS--QS--------------QSHMSQSVSSPSSMEVGVVPDG : 225
MtCOLc   : PVHGNFD----------------------PFVSAYKNNNVHLHTELETPS-----------------QSQISQSVSS-SSMDVGVVPDA : 214
PsCOLc   : PDHGNFD----------------------LFAYAFKNNNVQPHTEIETPSPSPS--------------QSQISHSVVS-SSMEVGVVPDG : 225
CaCOLc   : PVHGNFE----------------------PFS--YKINSVQLQTDLETPS-----------------QSPMSHSVSS-SSMDVGVVPDG : 212
CaCOLb   : PVQSHSK----------------------TAAEH----YSDINIDFSNSK--PF-------------TYSYNHNASS-PSMEVGVVPDG : 260
PsCOLb   : PVQSHSK----------------------TVTEH----YSDINNDFSTSK--PF-------------TYNYNHSVSS-PSLEVGVVPDG : 221
MtCOLb   : PVQSHSK----------------------TATEHEHEHYSDINIDFSNSK--PF-------------TYNFNHTVSS-PSMDVGVVPDG : 223
AtCOL3   : PVQNKLF----------------------LNEDYFNFDLS-ASK--ISQQG-----------------FNFINQTVST-RTIDVPLVPES : 207
AtCOL4   : PVENRTVRIP-------------------TVNENCFEMDFTGGSKGFTYGGG-----------------YNCISHSVSS-SSMEVGVVPDG : 306
GmCOL4a  : PVQTKPSLAPPPINNHQ----------HHHQSETCFDIDFCRSKLSSFNYP-----------------SQSLSQS-VSSSSLDVGVVPDG : 265
GmCOL4b  : PVQ-KPSLAPPLINNH----------HHHQSETCFDVDFCRSKLSSFNYP-----------------SNSLSQS-VSSSSLDVGVVPDG : 262
MtCOLd   : PVQTKPTPAP--MMN-------------HNSEGCFDIDFCRSKLSSFNYP-----------------SHSISHS-VSSSSLDVGVVPDG : 260
PsCOLd   : PVQTKPTPAP--MMNH------------HNSEACFDIDFCRSKLSSFNYP-----------------SHSISQS-VSSSSLDVGVVPDG : 262
CaCOLd   : PVQTKPTPAP--MMN-------------HNSESCFDIDFCRSKLSSFNYP-----------------AQSISQSQVSSSSLDVGVVPDG : 263
AtCOL5   : PVQTKTEPLP--LTN--------------NDHCFDIDFCRSKLSAFTYP-----------------SQSVSHS-VSTSSIEYGVVPDG : 258
AtCOL9   : GHGGIDSLFH-----------------------KHQTAPEGG-----NSVQPA----------------GSN----DSFMSSKTEPIICFA : 259
AtCOL10  : EHGGIGSLFE-----------------------KDE-AHEG------SMQQPA----------------LSNNASADSFMTCRTEPIICYS : 259
GmCOL8a  : ENGGINSLFEAKDMSAS---------AGDSHCQGAVAAEGSSAGLVNPIQPA----------------CSNAASADSVMSTKTEPIVCFT : 297
GmCOL8b  : ENGGINSLFETKDMSAS---------AGDSHCQGAVAAEGSSAELVNAIQPA----------------CSNAASADSMMSTKTEPIVCFT : 303
GmCOL9a  : ENGGIDSLFGTKDMSA-----------GDFSCEDAIAAEGSSVGQVNVMQPA----------------CSNAASADSILSTKTEPILCFT : 300
GmCOL9b  : ENGGIDSLFGTKGMSA-----------GDSNCQEAIAAEGSSVGQVNAMQPA----------------CSNAASADSILSTKTEPILCFT : 300
MtCOLe   : ENGGFNSLFGAKAMSA-----------GDSNCQDANAAEGSSIGHVNAAQPA----------------CSTAASADSILSTKTEPNLCIT : 300
PsCOLe   : ENGGFNSLFRAKAMSA-----------GDSNCQDAIVAEGSSIGHVNAAQPA----------------CSNAASADSILSTKTEPNFCIT : 315
CaCOLe   : ENGGFNSLFGAKEMSA-----------GDSDCQDAIAAEGSSVGHFNATQPA----------------CSNAASADSILSTKTEPNLCIT : 300
MtCOLf   : ENGGINSLFERKDMSAS---------AGDSNCQGAFAAEGSSARFVSAIQPE----------------CSNAASADSILSTKTEPVIYFT : 303
PsCOLf   : EKGGIDSLFERKDMSAS---------AGDSNCRGAIAA-------------------------------SSDSMLSTKTEPILYFT : 284
CaCOLf   : EKGGIDSLFETIDMYAS---------AGDLNCQGAVAAEGSSAGFVGAIQPA----------------CSNTASADSILSSKTEPILNFT : 303
AtCOL11  : ------CYEYEDSFKE------------ENNIGLPSLLLPTLSG-----------------------------------NVVPNMSLSMS : 236
AtCOL12  : DHTVPNCLLIDKTNTSS---------FTGSNFTVDKALEASPPGQQM-----------------------------NINTGLQLPLS : 257
GmCOL10a : EDGGMDCLLMDKNISVT---------ESNSLIESALEASSSIQQDCVAFQSSRAGGS--ASVMQVINSNTNSALMNPSCTRNISLGFPQG : 318
GmCOL10b : EDGGMDCLLMDKNIAVT---------ESSSLIESAMEASSSIQQDCVAFQSSRAGGS--ASVMQVINSNTNCALMNPSCTKNISLGFPQG : 320
MtCOLj   : EDGGIECLLMDKNIPVT---------KCSSHIETAVEASSSVQQDCMIFPSSGAGGS--TNLMQGFNNSANCALMPPSCNRSMPLEFPQS : 416
PsCOLj   : EDGGMKCLLMDKNIQVT---------KCTSLTEIAAEALSPLQQDCV------AGGS--TSVMQGINNNANSALMTPSSNSSITMGFPQS : 325
CaCOLj   : EDGGIDCLLMENNTKCS-----------SHIIETAVEASSSVQQDYVDYQSSGAS-------------NANCALMTPSCNQGMSLGFPQS : 348
GmCOL11a : EREPEADIFPSHE----------------------WHRESSEPMYQVVPPDPL----------------MRTYTEEIPFKHSTSAVGETQT : 270
GmCOL11b : ERDPEADIFPSHE----------------------WHRESSEPMHQVVPPDPS----------------MGTYTEEIPVKHSTSAVGENHT : 272
MtCOLg   : DCDVKADIVPSNE----------------------WLRESSEPMYQVVPVDTS----------------FKAHTEEISVKHSVSSVGEPHT : 280
PsCOLg   : EHDTKADIVPINE----------------------WLRESSEPIYQVVPVDTS----------------IRAHTEEFPVKHSVSSVGESHT : 281
CaCOLg   : EHDTKDDIIPSNE----------------------WLGESGEPVYQVVPVDTS----------------FKAHTKEIPVKHAVPSAEEPHS : 276
GmCOL12a : ERDVEANIFPSYEVGV----------------FCWHGESSDPRNQIVPSDTSP---------------LRDYGDIVSAEDGSFTI--TVT : 259
GmCOL12b : ERDVEASMFPSYEAGV----------------FCWHGESSDPTNQIVPSDTSS---------------LRDFGEVVSAEDGSFTIPGTGT : 267
AtCOL13  : QF-LAPDLFSTCELESG---------------LKWFDQQDHEDFPYCSLLKN----------------LSESDEKPENVDRESSVMVPVS : 256
AtCOL14  : GQSRNPDEPSPVETK---------------GSTFTFNNVTHLKNDTRTTNMNA---------------FKESYQQEDSVHSTSTKGQET- : 304
AtCOL15  : GQSRGPEDTSRVEAAYV---------GKGAASSFTINNFVDHMNETCSTNVKG---------------VKEIKK-DDYKRSTSGQVQPTK : 333
GmCOL13a : QKSRDCHE-PRVVTF------------DGLEVPKLFQDEHNMKYSTIGDDIDI---------------LSRNNQSDQSSSSHAKKKEENN : 280
GmCOL13b : QKSTDCNE-PRV----------------------DVHNMNYSTIGDDIDI---------------LSRNNQSDQSSSSHAKRKEESN : 259
MtCOLh   : QKSRDMTY-DGVEN-------------ASLSIPKSLQDVHNMNCSTLGD--DI---------------LSRNNQSDQSSSSHVKKKVESN : 287
PsCOLh   : QKSRDMTY-DGVE----------------VVPKSLQDVH---YSTLGD--DF---------------LSRSNHSDQSSSSHVKKKVDSN : 279
CaCOLh   : QKSRDMTY-DGAEE-------------ASLSVPKSMKHVLDMSSSTLGD--YI---------------LSRNNQSDQSSSSHVKEKVESN : 291
AtCOL6   : EEGVVTREVHDQD----------EGDETSPFEISFDYEYTHKTTFDEGEEDE----------------KEDVMKNVMEMGVN--------- : 274
AtCOL16  : EIKAMSMDIFDDDRK--------DVDGTVPFELSFDYESSHKTSEEEVMKNVES------SGECVVKVKEEEHKNVLMLRLNYDSVISTW : 306
GmCOL7a  : MEDQEESPLVEMEMD------MVVGRDDQSFELSFDY----ETCEEV---KVCDLGLGNELGAKKENDD-EVKKNKISLQLDYEAVIIAW : 311
GmCOL7b  : VEE-EESPLMEMDM----------GRDDQSFELSFDY----ETCEEVKEMKVSDLELGNELGEMKENDDDEVKRKKVSLQLDYEAIIIAW : 310
CaCOLk   : VEDEESHEVVVLEGYKIEGDNIEMGKES-SFELNFDYDDSNETSEEVKEKVELEIGK---CGEQNENNN-NKGKRKISLQLDYDAVIIAW : 338
MtCOLk   : EEGEECYEVVEGD------NMMEIGKES-SFELNFDYDDSHETCEEVKEK----------CGEQNENNDYNKGKRKISLQLDYDAVIIAW : 329
PsCOLk   : EEGGEIHEVVEGDD-----NTKDMGKES-SFELNFDY----ETCEEVKEKG----GL---DIEQDENKDCNKGKRKISLQLDYDAVIIAW : 336
GmCOL5a  : KVKDEEELDADDDTA-----CHLDSILDMN-GEAFNWNIVESESPAQAQEDEV---------------SKVGTKKDIFLRLNYEEVITAW : 323
GmCOL5b  : KVKDEEELDADDDTA-----CHLDSILDMN-SEAFNWNIAESESPAQAQEDEK---------------SLLLGQVKVLHGLQEPHLNLVT : 284
GmCOL6a  : KVKDEE--EIDGDVA-----CYLESVF----DDAFHWNNIESVLSDAREEKEGVVACDGGVGD-----EEGGTKRDIFLRLNYDEVITAW : 266
GmCOL6b  : KVKGEE--EIDGDVA-----CYLESVFDMISDDAFHWNNIESVVSDAREEKGCVVPCDGGVS------EQGGIKRDIFLRLNYDEVITAW : 272
MtCOLi   : KVKDEE--LDD-----------LESVFDMTSDDVFHWNIDNNDVSLAQQEKEYMPLSNSSVGYSESVITKEETKRERFLRLNYEEVITEW : 305
PsCOLi   : KVKDEE--LDD-----------LESVFDMTSDEVFHWNMD-HDESMVHQEKEFMPLSNSSVGVSESVITKEETKRERFLRLNYEDVISAW : 314
CaCOLi   : KVKDEE--LDD-----------LESVFDMTSDEVFQWNIDCNNETVAQEEKE------CVVG---SFDAKEETKRERFLRLNYEDVISAW : 292
AtCOL7   : EENKVGFEINCKDLKR------VKDEDEEEEEAKCENGGSKDSDREASNDKD--------------------RKTSLFLRLDYGAVISAW : 290
AtCOL8   : EEDKTDGAEACP-----------GQYLMSCKKDYDNVITVSEKTEEIEDCY--------------------ENNARHRLNYENVIAAW : 239
```

Sequence alignment for Figure 3.1. Continued


AtCO : TAC-VTTASHPRTPKGT-------------------------VEQQPDPASQMITVTQLSPMDREARVLRYREKRKTRKFEKTIRYAS : 330
AtCOL2 : TAS-DITVQHPRTTKET-------------------------IDQLSGPPTQVV--QQLTPMEREARVLRYREKKKTRKFDKTIRYAS : 302
AtCOL1 : TTS-DATVSNPRSPKAV-------------------------TDQPPYPPAQML-----SPRDREARVLRYREKKKMRKFEKTIRYAS : 310
GmCOL1a : PMR-DVSIGHTRTPKGT-------------------------IDLFSGPPIQVP--SHFSPMDREARVLRYREKKKMRKFEKTIRYAS : 303
GmCOL1b : PMR-DVSIAHTRPPKGT-------------------------IDLFSGPPIQVP--SHFSPMDREARVLRYREKKKMRKFEKTIRYAS : 307
MtCOLa : ------TMTYSRPPKGT-------------------------IDLFSGPSIQMS--SHFSPMDREARVLRYREKKKTRKFEKTIRYAS : 351
PsCOLa : TMRDATTMSYSRPSKGT-------------------------IDLFSAPPIQMT--SHFSPMDREARVLRYLEKKKTRKFEKTIRYAS : 387
CaCOLa : ------TMTYSRPPKGT-------------------------IDLFSGPPIQMT--SHFSPMDREARVLRYREKKKTRKFEKTIRYAS : 332
GmCOL2a : TIS-DISMSHSKSPIGT-------------------------TDLF--PPLPMP--SHLTPMDREARVLRYREKKKTRKFEKKIRYAS : 316
GmCOL2b : TVS-GISMSHSKSPIGT-------------------------NDLF--PPLLMP--SHLTPMDREARVLRYREKKKTRKFEKKIRYAS : 321
GmCOL3a : NTMSETSNCSYSKVPPV-------------------------T----------VTVTAQFSAADREARVLRYREKRKNRKFEKTIRYAS : 276
GmCOL3b : NTMSEISNCSYSKVAP-------------------------------------VTVTAQFSAADREARVLRYREKRKNRKFEKTIRYAS : 277
MtCOLc : NTVPEISN--CGYGT---VAVDREARVMRYREKRKNRKFEKTIRYAS : 256
PsCOLc : EAVSEISNGGCGKVV---VAADREAKVMRYREKRKNRKFEKTIRYAS : 269
CaCOLc : NAVSDISN--CGYGKA---VAADREARVLRYREKRKNRKFEKTIRYAS : 255
CaCOLb : NAMSEISYCCYGRAA-------------------------------------TEAVQITAADREARVLRYREKRKNRKFEKTIRYAS : 310
PsCOLb : NVMSEMSYCGYGR-------------------------------------TEAVQITAAGREARVMRYREKRKNRKFEKTIRYAS : 269
MtCOLb : NVMTEISYCSYQNNGDC-------------------------SYQTTATETAPMTVAVPMTAVEREARVMRYREKRKNRKFEKTIRYAS : 287
AtCOL3 : GGVT-------AEMTNT-------------------------ETP-----------AVQLSPAEREARVLRYREKRKNRKFEKTIRYAS : 253
AtCOL4 : GSVADVSYPYGGPATSG-------------------------ADPG-------TQRAVPLTSAEREARVMRYREKRKNRKFEKTIRYAS : 363
GmCOL4a : NTVSDMSYSSG-------------------------------IVVSG---GQGATQLCGMDREARVLRYREKRKNRKFEKTIRYAS : 317
GmCOL4b : NTVSDMSYSFGRNSSDS-------------------------SGIVVVSGNSVGQGATQLCGMDREARVLRYREKRKNRKFEKTIRYAS : 326
MtCOLd : NTVSEISYNFG---SES-------------------------MVSGGVNSSNQGVQGATQLCGMDREARVMRYREKRKNRKFEKTIRYAS : 322
PsCOLd : NVVSEFSYTFG---SES-------------------------MVSGGVSSNNQGVQGATQLCGIDREARVLRYREKRKNRKFEKTIRYAS : 324
CaCOLd : NAVPEMSYSFGRNSSESS-----------------------GMVSGGVNS-QQGVQVATQLCGMDREARVLRYREKRKNRKFEKTIRYAS : 329
AtCOL5 : NTNNSVN-------------------------------------RSTITSSTTGGDHQASSMDREARVLRYREKRKNRKFEKTIRYAS : 309
AtCOL9 : SKPAHSNISFSGVTGE-SSAGDFQECGA--------SSSIQLSGEP-PWYPPTLQDNNACSHSVTRNNAVMRYKEKKKARKFDKRVRYAS : 339
AtCOL10 : SKPAHSNISFSGITGE-SNAGDFQDCGA--------SSMKQLSREPQPWCHPTAQDIIASSHATTRNNAVMRYKEKKKARKFDKRVRYVS : 340
GmCOL8a : ARQSQSNISFSGVTKD-S-AGDYQDCG---------ASSMLLMGEP-PWCPPCPES---SLHSANRSNAVMRYKEKKKTRKFEKKVRYAS : 372
GmCOL8b : ARQSLSNISFSGVTKD-S-VGDYQDCG---------ASSMLLMGEP-PWCPPCPES---SLHSANRSNAVMRYKEKKKTRKFEKKVRYAS : 378
GmCOL9a : GRQAQSNLSFSGVTGD-SSAGDYQDCG---------ASSMLLMGEP-PWFAPCPEN---SLQSANRSNAVMRYKEKKKTRMFDKKVRYAS : 376
GmCOL9b : GRQTQSNLSFSGVTGE-SSAGDYQDCG---------ASSMLLMGEP-PWFAPCPEN---SLQSANRSNAVMRYKEKKKTRMFDKKVRYAS : 376
MtCOLe : AKQSQSSLSFSGINED-GGAGDYQDCG---------ASSMLLMGEP-PWLNTCPENE-LQLQSANRCSAVMRYKEKKKTRKFDKRVRYAS : 378
PsCOLe : TKQSQSSLSFSGINED-GCAGDYQDCG---------ASSMFLMGEP-PWLNTCPENE-LQLQSANRCSAVMRYKEKKKTRKFDKRVRYAS : 393
CaCOLe : AKQAQSNLSFSGINEESGGAGDYQDCG---------ASSMLLMGEP-PWLNTCPEND-LQLQSANRSSAVMRYKEKKKTRKFDKRVRYAS : 379
MtCOLf : ERQS--NLSFSGVNKD-ASAGDYQECG---------TSSMLLTGEP-PWCPPCPENS---IQSANRSNAVMRYKEKKKNRKFDKKVRYAS : 377
PsCOLf : EMQSQSNVSFSGVIND-ASAGDNQECGS--------ASSMLLTGEP-PWCPPCLESSS--IQSANRSNAVMRYKEKKKNRKFDKKVRYAS : 362
CaCOLf : ARQSQSNISFSGGTKD-ASAGNSQDCG---------GSAMHVSGEP-PWRVPFAESS---TQSANRSNAVMRYKEKKKNRKFDKKVRYAS : 379
AtCOL11 : -----------NLT-GESNATDYQDCGI--------SPGFLIGDSP--WESNVEVSFN----PKLREEAKKRYKQKKSKRMFGKQIRYAS : 300
AtCOL12 : PVLFGQIHPSLNIT-GENNAADYQDCGM--------SPGFIMSEAP--WETNFEVSC-----PQARNEAKLRYKEKKLKRSFGKQIRYAS : 331
GmCOL10a : -VHSKMPLQFPNIV-GENNSTEYQDCRF--------SQVFLPGESP--WESNLEGTC-----PQARDKAKMRYNEKKKTRTFGKQIRYAS : 391
GmCOL10b : LVHSNMPLQFPNIV-GENNSTEYQDCGL--------SRV---GESP--WESNLEGTC-----PQARDKAKMRYNEKKKTRTFGKQIRYAS : 391
MtCOLj : QTHSGISIQLPNIN-GESNVAELLDCGL--------PPVFHPGESH--WESNLEGAC-----PQARDKAKMRYQEKKKTRTFGKQIRYAS : 490
PsCOLj : QIHSGTSIELPNLN-GENNVTELLNCEL--------PPVFHPGESP--WEPNLEGTC-----AEAREQAKMRYQEKKKTRTFGKQIRYAS : 399
CaCOLj : QIHSDISIQLPTMISGENMPKEFHDG----------------GLPP--WESNLEGKC-----PLAREKATMRYMEKKKTRTFDKQIRYAS : 415
GmCOL11a : YGDNGG-----------------------------KPSISLKSET-LSTTPKAA--ACELTSQERDSALLRYKQKKKTRRFDKHIRYES : 327
GmCOL11b : YGDNEG-----------------------------KPSISLKSET-LSTTPKAA--ACELTSQERDSALLRYKQKKKTRRFDKHIRYES : 329
MtCOLg : HCNNGGTPSEPLNHCN---------NGG--------TPSEYVKSET-LSTTSKAVPPPYELASQERDSALLRYKQKKKTRRYDKHIRYES : 352
PsCOLg : HM-----------------------------------TPSESIKSEA-LSTTFKPLPPPYELASQERDSALLRYKQKKKTRRYDKHIRYES : 336
CaCOLg : HSNNGG-----------------------------TPSESLKSET-LSTTPRAVPAPYELASQERDSALLRYKQKKKTRRYDKHIRYES : 335
GmCOL12a : HANFNNQ--------------------G--------KPSNSFNAEN-LSPTPKAT--PYELTSHERDSALLRYREKKKSRRYDKHIRYES : 318
GmCOL12b : QANFNNE--------------------G--------KPSNSFNSEN-LSPTPKAT--PYELTSHERDSALLRYREKKKTRRYDKHIRYES : 326
AtCOL13 : GCLNR-----------------------------CEEETVMVPV-ITSTRSMT---HEINSLERNSALSRYKEKKKSRRYEKHIRYES : 311
AtCOL14 : SKSNN----IPAAIHSHKSSNDSCGLHC--------TEHIAITSNR-ATRLVAVTNADLEQMAQNRDNAVQRYKEKKKTRRYDKTIRYET : 381
AtCOL15 : SESNN----RPITFGSEKGSNSSSDLHF--------TEHIAGTSCK-TTRLVATK-ADLERLAQNRGDAVQRYKEKKKTRRYDKTIRYES : 409
GmCOL13a : KKAKGGLSSESKLFESIPYNG-TNNVVV--------MEHLVGGNEN-VGTLTARV--SLEELAKNRGDAVLRYKEKKKTRRYDKHIRYES : 358
GmCOL13b : KKARGGLSSESTLFESIPYSG-TNNVVV--------MEHLVGGNEN-VSTLKARV--SLQELAKNRGDAVLRYKEKKKTRRYDKHIRYES : 337
MtCOLh : KKTRDGLSSESKLIESITYSG-ADSVPV--------MEHLLSGSEN-VSNINAKI--SLEEHTRNRGDAVLRYKEKKKTRRFDKHIRYES : 365
PsCOLh : KKTRDGLPTESKLVESITYNGAADNVPV--------MEYLLSGSEN-VSNINGKV--SLEEQVRNRGDAVLRYKEKKKTRRFDKHIRYES : 358
CaCOLh : KKTKDGLSSQSKLTKSITHGT-NNIVPV--------MEPLLSGSEN-DSNIKAKV--ILEEQARNRGDAVLRYKEKKKTRRFDKHIRYES : 369
AtCOL6 : --------------------EMSGGI---

Sequence alignment for Figure 3.1. Continued

```
                  640         *       660         *       680
AtCO     : RKAYAEIRPRVNGREAKR-EIEAE-EQGFNT----------MLMYNTGYGIVPSF- : 373
AtCOL2   : RKAYAEIRPRIKGREAKRIETEAEAEEIFST----------SLMSETGYGIVPSF- : 347
AtCOL1   : RKAYAEKRPRIKGREAKKDVDEEANQAFST----------MITFDTGYGIVPSF- : 355
GmCOL1a  : RKAYAETRPRIKGREAKRTDVEAEVDQMFST----------TLITEVGYGIVPSF- : 348
GmCOL1b  : RKAYAETRPRIKGREAKRTDVEAEVDQMFST----------TLITEVGYGIVPSF- : 352
MtCOLa   : RKAYAETRPRIKGREAKRTDVEAEVDQMFST----------SLITEVGYGIVPSF- : 396
PsCOLa   : RKAYAETRPRIKGREAKRTDVEAEVDQMFST----------TLITEVGYGIVPSFV : 433
CaCOLa   : RKAYAETRPRIKGREAKRTDVEAEVDQMFSS----------TLITEIGYGIVPSF- : 377
GmCOL2a  : RKAYAETRPRIKGREAKRTDVEAEVDQMFST----------TLFTEVGGSIFPTF- : 361
GmCOL2b  : RKAYAETRPRIKGREAKRTDVEAEVDQMFST----------KLFNEVGGSIFPTF- : 366
GmCOL3a  : RKAYAEARPRIKGREAKR---------TDP-------------DPLAGYGVVPSC- : 309
GmCOL3b  : RKAYAETRPRIKGREAKR---------TDA-------------DPLAGYGVVPSC- : 310
MtCOLc   : RKAYAETRPRIKGREAKR---------TDAV------------DSISGYGVVPTC- : 290
PsCOLc   : RKAYAETRPRIKGREAKR---------TDAV------------DSLGGYGVVPTC- : 303
CaCOLc   : RKAYAETRPRIKGREAKR---------TDAV------------DSLAGYGVVPTC- : 289
CaCOLb   : RKAYAETRPRIKGREAKRTDMNVNVNLIAED------------ESYTGYGVVPSC- : 353
PsCOLb   : RKAYAETRPRIKGREAKRTDLNMNVNLIGED------------ESYDGYGVVPSC- : 312
MtCOLb   : RKAYAETRPRIKGREAKRSDLNMN--LIAED------------E----YGVVPSC- : 324
AtCOL3   : RKAYAEMRPRIKGREAKRTDSREND--GGDV------------GVYGGFGVVPSF- : 294
AtCOL4   : RKAYAEMRPRIKGREAKRTDTNESNDVVGHG------------GIFSGFGLVPTF- : 406
GmCOL4a  : RKAYAETRPRIKGREAKRTEIDSDVERLYSP-------GAAALMLDTPYGVVPTF- : 365
GmCOL4b  : RKAYAETRPRIKGREAKRTEIDSDVERLYSP-------GPAVLMLDTPYGVVPSF- : 374
MtCOLd   : RKAYAETRPRIKGREAKRTEIDSDVDRLYN--PADPLSVPSSMLMDCPYGVVPTF- : 375
PsCOLd   : RKAYAETRPRIKGREAKRTEIDSDIDGLYNN-PADPLTIPASLLIDSPYGVVPTF- : 378
CaCOLd   : RKAYAETRPRIKGREAKRTEIDSEVDRLYDPIEIDHLTVPSSLLIDSPYGVVPTF- : 384
AtCOL5   : RKAYAESRPRIKGREAKRTETEND-DIFLSH--------VYASAAHAQYGVVPTF- : 355
AtCOL9   : RKARADVRRRVKGREVKAGEAYDYDPLTPTRSY----------------------- : 372
AtCOL10  : RKERADVRRRVKGREVKSGEAYDYDPMSPTRSY----------------------- : 373
GmCOL8a  : RKARADVRKRVKGREVKAGDVYDYDPLNQTRSY----------------------- : 405
GmCOL8b  : RKARADVRKRVKGREVKAGDVYDYDPLNQTRSC----------------------- : 411
GmCOL9a  : RKARADVRRRVKGREVKAGDVYDYDPLSTTRSF*---------------------- : 409
GmCOL9b  : RKARADVRRRVKGREVKAGDVYDYDPLSTTRSC----------------------- : 409
MtCOLe   : RKERADVRRRVKGREVKAGEAYDYDPLSQTRSY----------------------- : 411
PsCOLe   : RKERADVRRRVKGREVKAGEAYDYDPLSQTRSY----------------------- : 426
CaCOLe   : RKERADVRRRVKGREVKAGDVYDYDPLSQTRSY----------------------- : 412
MtCOLf   : RKARADVRKRVKGREVKAGETYDYDPLSQTRSC----------------------- : 410
PsCOLf   : RKARADVRKRVKGREVKAGETFDYDPLSETRSF----------------------- : 395
CaCOLf   : RKARADVRKRVKGREVKAGESYDYDPLNETRSF----------------------- : 412
AtCOL11  : RKARADTRKRVKGREVKSGETFEYDPSLVM------------------------- : 330
AtCOL12  : RKARADTRKRVKGREVKAGDSYDYDPSSPTTNN--------------------- : 364
GmCOL10a : RKARADTRKRVKGREVKAGEAYDYDPLGTRDI---------------------- : 423
GmCOL10b : RKARADTRKRVKGREVKAGEAYDYDPLGTRDI---------------------- : 423
MtCOLj   : RKARADTRKRVKGREVKAGEAYDYDPLLSDH--------------------- : 521
PsCOLj   : RKARADTRKRVKGREVKAGEAYDYDPLLSDY--------------------- : 430
CaCOLj   : RKARADTRKRVKGREVKAGEAYDYDPLLSDS--------------------- : 446
GmCOL11a : RKVRAESRVRVKGREAKMEHEH------------------------------ : 349
GmCOL11b : RKVRAESRVRVKGREAKMGHEH------------------------------ : 351
MtCOLg   : RKVRAESRTRVKGREAKIDH-------------------------------- : 372
PsCOLg   : RKVRAETRTRVKGRLEEENKKVCEPNDTRS--------------------- : 366
CaCOLg   : RKVRAESRTRVRGREAKIEH-------------------------------- : 355
GmCOL12a : RKVRAESRMRIKGREVRDETQK------------------------------ : 340
GmCOL12b : RKVRAESRMRIKGREVKDETQK------------------------------ : 348
AtCOL13  : RKVRAESRTRIRGREAKAADP-------------------------------- : 332
AtCOL14  : RKARAETRLRVKGREVKAT---DP---------------------------- : 402
AtCOL15  : RKARADTRLRVRGREVKAS---EAPYP------------------------- : 433
GmCOL13a : RKARADTRKRVRGREVKAS---DVQA*------------------------- : 381
GmCOL13b : RKARADTRKRVRGREVKAS---DVQA-------------------------- : 360
MtCOLh   : RKARADTRKRVRGREVKAT--DDIQEG------------------------- : 390
PsCOLh   : RKARADTRKRVRGREVKAGD-DDIQEG------------------------- : 384
CaCOLh   : RKARADTRKRVRGREVKATDTSDIQAG------------------------- : 396
AtCOL6   : ------------------------------------------------------ :   -
AtCOL16  : RKLNAEKRPRMKGREVKRASLAAAA----------SPLGVNY---------- : 417
GmCOL7a  : RKLNAEKRPRMKGREVKRASFAPPT----------FPLLNK---------- : 419
GmCOL7b  : RKLNAEKRPRMKGREVKRASFAPPT----------FPLLNK---------- : 418
CaCOLk   : RKLNAEKRPRMKGREVKRASFAVPT----------FPLLK----------- : 445
MtCOLk   : RKLNAEKRPRMKGREVKRASFAVPT----------FPLLK----------- : 436
PsCOLk   : RKLNAEKRPRMKGREVKRTSFAVPT----------FPLLK----------- : 451
GmCOL5a  : RKLNAEKRPRMKGREVKRTPFVG---------ATALPA------------- : 427
GmCOL5b  : RKLNAEKRPRMKGREVKSVPLLEPLLCQLNYQSSPLT--------------- : 392
GmCOL6a  : RKLNAEKRPRMKGREVKR-TCFVGA--------NAFPAYH---------- : 365
GmCOL6b  : RKLNAEKRPRMKGREVKR-TCFVGA--------NAFPAYH---------- : 371
MtCOLi   : RKLNAEKRPRMKGREVKRTACFAGG-------ATSFPTNYH---------- : 416
PsCOLi   : RKLNAEKRPRMKGREVKRTTCFAGG-------ATPFPTNYH---------- : 425
CaCOLi   : RKLNAEKRPRMKGREVKRTTCFAGTT------TTSFPTYQH---------- : 405
AtCOL7   : RKLNAEQRPRIKGREVKRTSLLT------------------------------ : 392
AtCOL8   : RKVNADKRPRMKGREVRRSLAIDS---------------------------- : 319
```

Appendix 3.20 Alignment of the 3 TFL1c proteins found in *Cicer arietinum* (Ca) with those from *Medicago truncatula* (Mt) and *Pisum sativum* (Ps), and identity matrix (%) derived from it. Alignment was performed in Genious 8 using ClustalW.

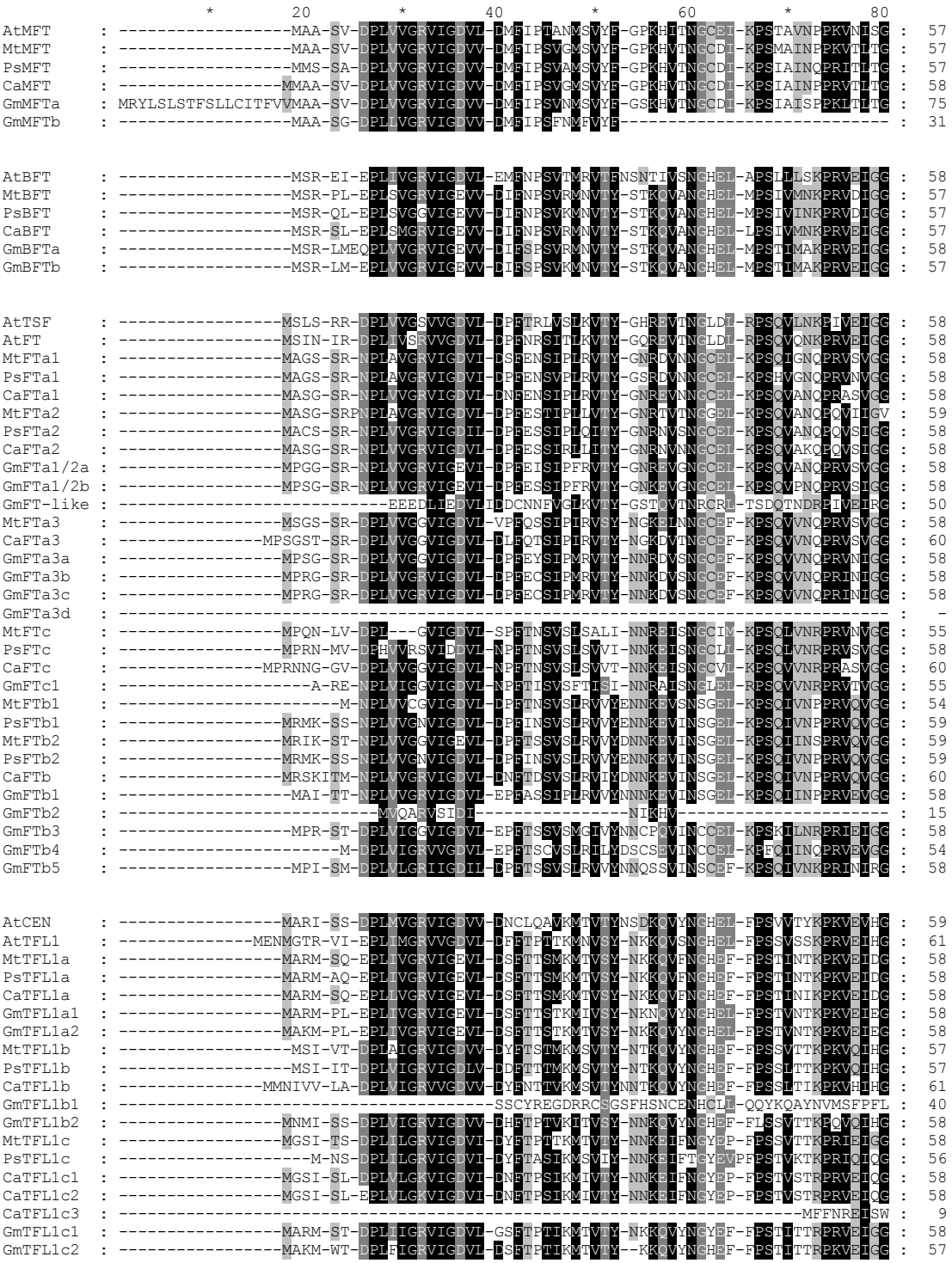
```
                    *         20         *         40         *
CaTFL1c1  :  MGSISLDPLVLGKVIGDVIDNFTPSIKMIVTYNNKEIFNGYE-PFPSTVS  :  49
CaTFL1c2  :  MGSISLEPLVLGKVIGDVIDNFTPSIKMIVTYNNKEIFNGYE-PFPSTVS  :  49
CaTFL1c3  :  ------------------------------------MFFN-----------  :   4
MtTFL1c   :  MGSITSDPLILGRVIGDVIDYFTPTTKMTVTYNNKEIFNGYE-PFPSSVT  :  49
PsTFL1c   :  ---MNSDPLILGRVIGDVIDYFTASIKMSVIYNNKEIFTGYEVPFPSTVK  :  47

                   60         *         80         *        100
CaTFL1c1  :  TRPRVEIQGGDMRSLFTLIMIDPDVPGPSDPYMREHLHWMVTDIPGTTDS  :  99
CaTFL1c2  :  TRPRVEIQGGDMRSLFTLIMIDPDVPGPSDPYMREHLHWMVTDIPGTTDS  :  99
CaTFL1c3  :  -R---EISWNFLFIFFLQIMIDPDVPGPSDPYMREHLHWMVTDIPGTTDS  :  50
MtTFL1c   :  TKPRIEIGGVDMRSLFTLIMIDPDVPGPSDPYMKEHLHWMVTDIPGTTDS  :  99
PsTFL1c   :  TKPRIQIQGGDMRSLFTLIMIDPDVPGPSDPYMKEHLHWMVTDIPGTTDS  :  97

                    *        120         *        140         *
CaTFL1c1  :  TFGKELTSYEIPKPNIGIHRYVFVLFKQKK--KHSITTPSSRDHFNTRSF  : 147
CaTFL1c2  :  TFGKELTSYEIPKPNIGIHRYVFVLFKQKK--KHSITTPSSRDHFNTRSF  : 147
CaTFL1c3  :  TFGKELTSYEIPKPNIGIHRYVFVLFKQKK--KHSITTPSSRDHFNTRSF  :  98
MtTFL1c   :  TFGKELTSYEKPKPNIGIHRYVFVLFKQEKGKKHSIVAPFSRDHFNTRAF  : 149
PsTFL1c   :  TFGKELTSYEKPKPNIGIHRYVFVLFKQKRGNKYSITCPFSRDHFNTRNF  : 147

                   160         *
CaTFL1c1  :  SMQNDLGVPVAAAYFNARRPTAGRKPTYI  : 176
CaTFL1c2  :  SMQNDLGVPVAAAYFNARRPTAARKPTYI  : 176
CaTFL1c3  :  SMQNDLGVPVAAAYFNARRPTAGRKPTYI  : 127
MtTFL1c   :  SAQNDLGVPVAAAYFNARRATAPRRRAS-  : 177
PsTFL1c   :  ADQNDLGVPVAAAYFNARRATAPRRR---  : 173
```

	CaTFL1c1	CaTFL1c2	CaTFL1c3	MtTFL1c	PsTFL1c
CaTFL1c1		98.9	81.6	83.1	82.7
CaTFL1c2	98.9		80.9	82.5	82.1
CaTFL1c3	81.6	80.9		69.7	68.1
MtTFL1c	83.1	82.5	69.7		86.7
PsTFL1c	82.7	82.1	68.1	86.7	

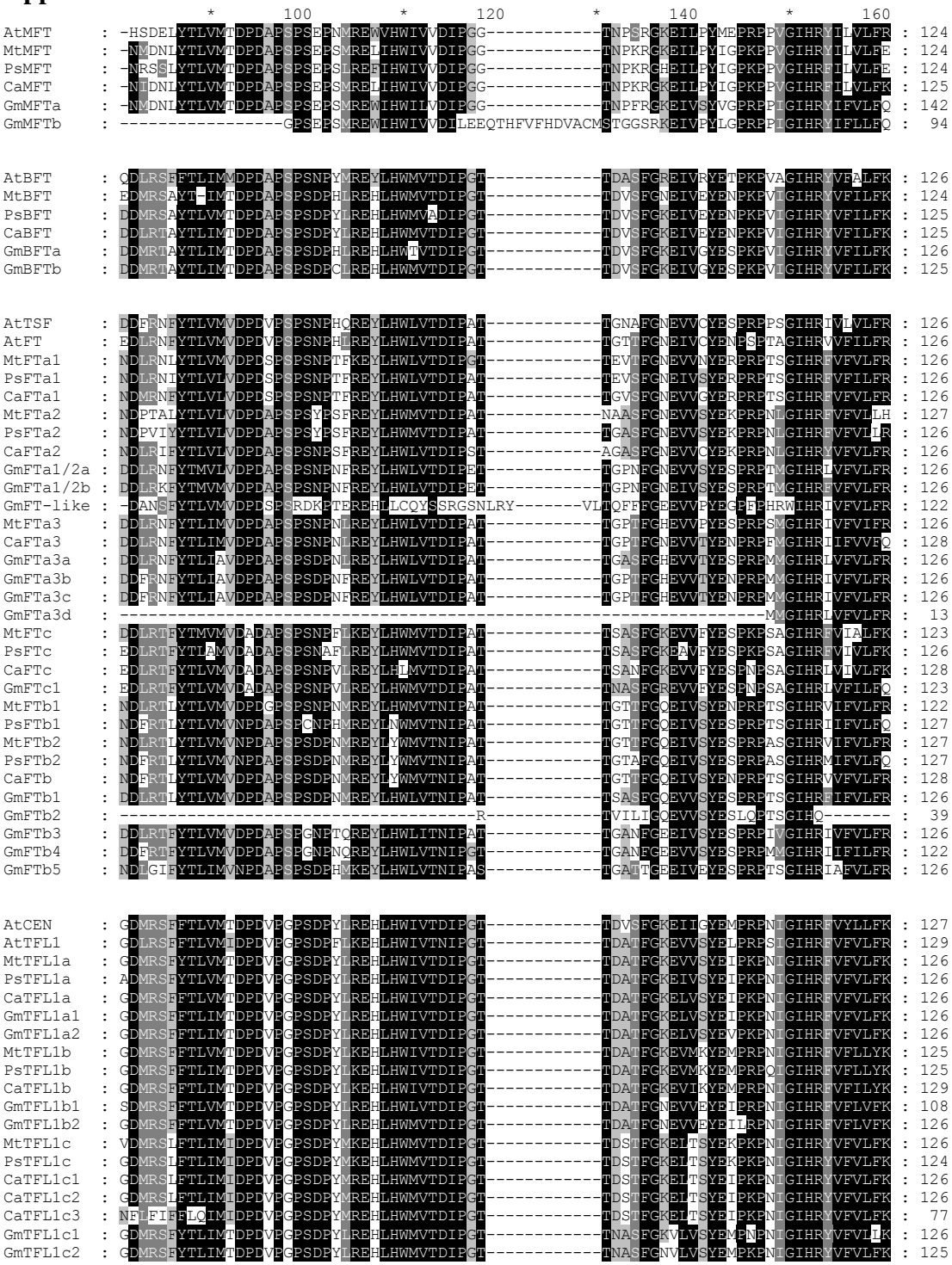
Appendix 3.21 Accession number of Phosphatidylethanolamine-binding protein (PEBP) sequences used in the alignment (appendix 3.22) and tree construction (Fig 3.2) in five plant species; *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt), *Pisum sativum* (Ps), *Glycine max* (Gm) and *Cicer arietinum* (Ca). Soybean accessions were obtained from Wang et al. (2015) and the protein sequences retrieved from the Soybean Knowledge Database [http://soykb.org; Joshi et al. (2012)]

Arabidopsis		**Medicago**		**Pea**	**Chickpea**		**Soybean**			
MFT	AT1G18100	Medtr8g106840		PsCam040701	LOC101504081		**GmMFTa**	Glyma05g34030	**GmMFTb**	Glyma08g05650
BFT	AT5G62040	Medtr0020s0120		PsCam044479	LOC101507903		**GmBFTa**	Glyma09g26550	**GmBFTb**	Glyma16g32080
CEN/ATC	AT2G27550	-		-	-		-		-	
TFL1	AT5G03840	**TFL1a**	Medtr7g104460	AY340579	**TFL1a**	LOC101506075	**GmTFL1b1**	Glyma12g30940	**GmTFL1c1**	Glyma10g08340
		TFL1b	Medtr2g086270	AY340580	**TFL1b**	LOC101508699	**GmTFL1b2**	Glyma13g39360	**GmTFL1c2**	Glyma13g22030
		TFL1c	Medtr1g060190	AY343326	**TFL1c1**	LOC101495644	**GmTFL1a1**	Glyma03g35250		
					TFL1c2	LOC101491943	**GmTFL1a2**	Glyma19g37890		
					TFL1c3	LOC101492277				
FT	AT1G65480	**FTa1**	Medtr7g084970	HQ538822	**FTa1**	LOC101497376	**GmFTc1**	Glyma19g28400	**GmFTb1**	Glyma08g47820
TSF	AT4G20370	**FTa2**	Medtr7g085020	HQ538821	**FTa2**	LOC101496618	**GmFTa3a**	Glyma16g26660	**GmFTb2**	Glyma18g53670
		FTa3	Medtr6g033040	-	**FTa3**	LOC101515383	**GmFTa3b**	Glyma16g04830	**GmFTb3**	Glyma18g53680
		FTb1	Medtr7g006630	HQ538824	**FTb**	LOC101505276	**GmFTa3c**	Glyma16g26690	**GmFTb4**	Glyma18g53690
		FTb2	Medtr7g006690	HQ538825			**GmFTa3d**	Glyma02g07650	**GmFTb5**	Glyma08g47810
		FTc	Medtr7g085040	HQ538826	**FTc**	LOC101508200	**GmFTa1/2a**	Glyma19g28390	**GmFT-like**	Glyma08g28470
							GmFTa1/2b	Glyma16g04840		

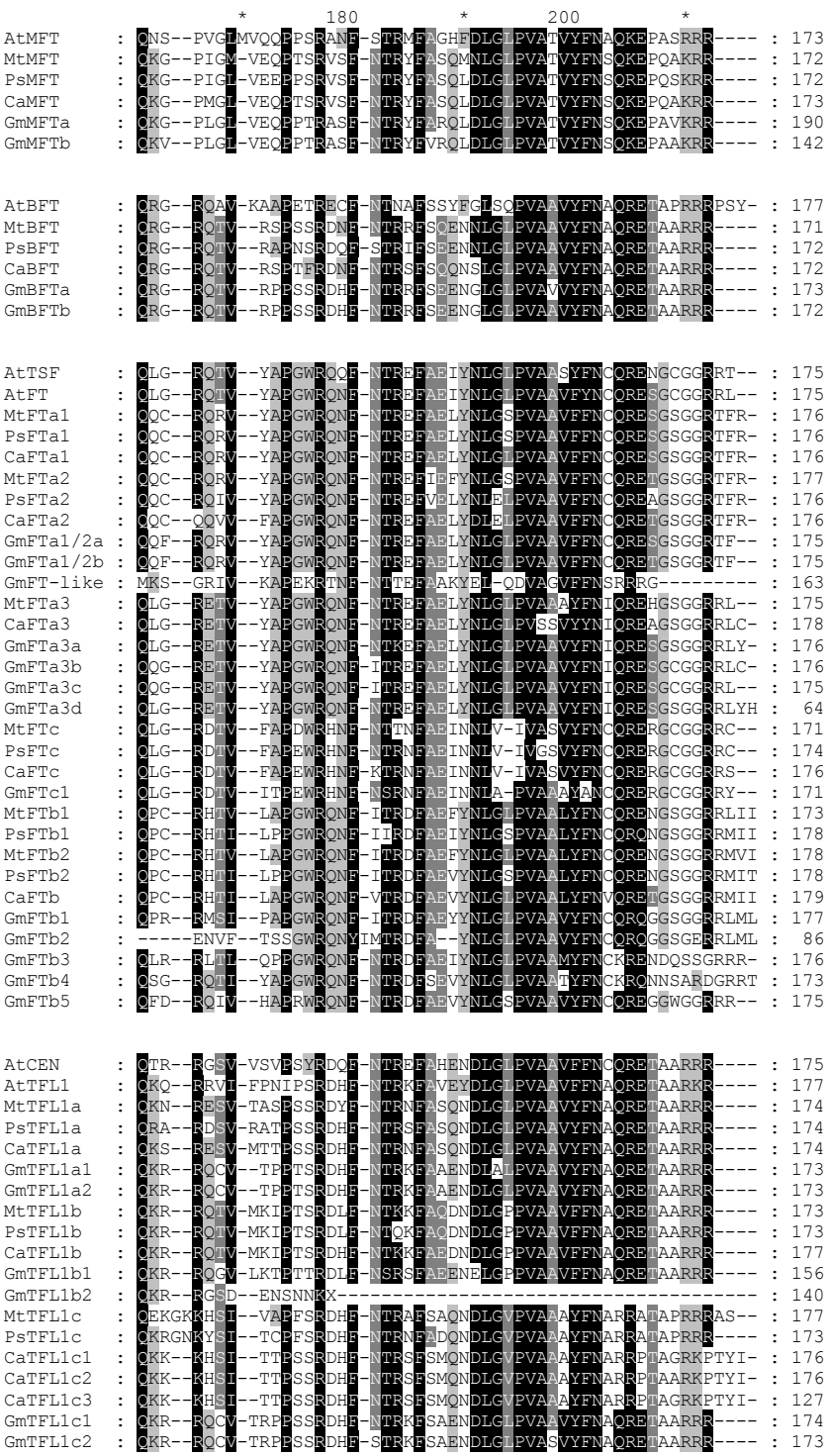
Appendix 3.22 Multiple sequence alignment of the PEBP sequences from five species described in appendix 3.21. Proteins were aligned in Geneious 8 using MAFFT. Residues conserved in more than 80% of the sites are presented with black background/while letters. Those with a conservation between 60-80% show grey background/white letters and those between 40-60% with grey background/black letters.



Appendix 3.22 Continued



Appendix 3.22 Continued



Appendix 3.23 Identity matrix (%) of MFT and BFT proteins in *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt), *Pisum sativum* (Ps), *Glycine max* (Gm) and *Cicer arietinum* (Ca), derived from the previous alignment (appendix 3.22).

	AtMFT	MtMFT	PsMFT	CaMFT	GmMFTa	GmMFTb	AtBFT	MtBFT	PsBFT	CaBFT	GmBFTa
MtMFT	76.3										
PsMFT	71.7	87.2									
CaMFT	76.3	94.8	89.5								
GmMFTa	75.7	85.5	80.8	85.5							
GmMFTb	48.6	52.2	50.0	53.3	59.2						
AtBFT	49.1	48.9	48.9	49.4	49.4	34.4					
MtBFT	51.1	53.2	53.2	53.2	52.0	35.9	69.0				
PsBFT	52.3	54.3	53.8	54.3	53.2	35.1	68.4	90.7			
CaBFT	51.7	53.8	53.2	54.9	53.2	36.8	70.1	91.3	88.4		
GmBFTa	52.0	53.4	52.3	53.4	52.9	36.0	67.4	87.3	84.4	85.5	
GmBFTb	52.3	53.8	52.6	53.8	53.2	36.2	69.5	88.4	87.2	87.2	96.5

Appendix 3.24 Identity matrix (%) of TFL1-related proteins in *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt), *Pisum sativum* (Ps), *Glycine max* (Gm) and *Cicer arietinum* (Ca), derived from the previous alignment (appendix 3.22).

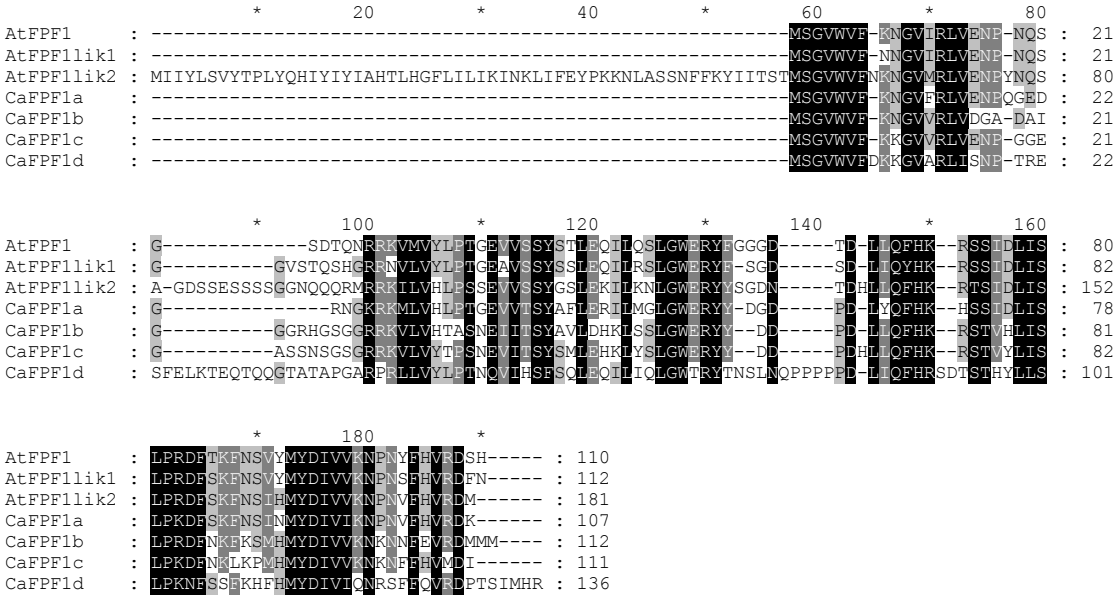
			TFL1a					TFL1b					TFL1c					
	AtCEN	AtTFL1	Mt	Ps	Ca	Gm-a1	Gm-a2	Mt	Ps	Ca	Gm-b1	Gm-b2	Mt	Ps	Ca-c1	Ca-c2	Ca-c3	Gm-c1
AtTFL1	66.3																	
MtTFL1a	73.1	73.0																
PsTFL1a	72.6	71.8	94.3															
CaTFL1a	73.7	73.0	94.8	93.1														
GmTFL1a1	72.6	74.7	86.2	85.6	87.4													
GmTFL1a2	73.7	75.9	86.8	86.2	87.9	97.1												
MtTFL1b	74.7	72.3	74.6	72.8	74.6	75.1	76.3											
PsTFL1b	72.4	69.9	73.4	71.7	73.4	74.6	75.7	94.8										
CaTFL1b	73.1	69.5	71.4	70.3	73.1	73.7	74.9	89.1	89.1									
GmTFL1b1	60.3	55.8	60.9	60.9	61.5	62.8	62.8	62.8	61.5	61.5								
GmTFL1b2	70.5	66.7	70.2	67.4	69.5	68.6	69.3	75.7	73.6	70.5	57.8							
MtTFL1c	62.1	65.3	67	67.6	67.6	68.0	69.7	66.9	65.1	63.8	48.1	64.1						
PsTFL1c	60.6	63.2	69.5	68.4	70.7	70.5	71.7	66.7	64.9	64.6	49.1	62.2	86.1					
CaTFL1c1	62.3	64.4	69.5	69.5	71.8	70.5	70.5	64.2	63.6	61.7	50.6	65.0	83.1	82.1				
CaTFL1c2	62.3	65.5	70.7	70.7	73.0	71.7	71.7	64.2	63.6	61.7	51.3	64.3	82.5	81.5	98.9			
CaTFL1c3	58.4	61.6	67.2	68.8	71.2	68.5	68.5	58.4	58.4	60.0	58.4	57.2	75.8	75.4	88.2	87.4		
GmTFL1c1	70.9	69.0	79.9	78.2	78.7	81.0	82.2	74.0	73.4	72.0	56.4	67.4	70.5	68.4	71.3	71.3	65.6	
GmTFL1c2	70.3	67.2	78.7	77.6	78.2	80.5	82.8	73.4	72.8	71.4	56.4	68.8	68.2	66.1	69.0	69.0	64.8	93.1

Appendix 3.25 Identity matrix (%) of the FT-related proteins in *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt), *Pisum sativum* (Ps), *Glycine max* (Gm) and *Cicer arietinum* (Ca), derived from the previous alignment (appendix 3.22).

			FTa1/a2								FTa3						FTc				FTb									
	AtFT	AtTSF	Mt-a1	Ps-a1	Ca-a1	Mt-a2	Ps-a2	Ca-a2	Gm-a	Gm-b	Mt	Ca	Gm-3a	Gm3-b	Gm-3c	Gm3-d	Mt	Ps	Ca	Gm	Mt-b1	Mt-b2	Ps-b1	Ps-b2	Ca	Gm-b1	Gm-b2	Gm-b3	Gm-b4	Gm-b5
AtTSF	81.7																													
MtFTa1	71.4	68.6																												
PsFTa1	72.6	69.1	92																											
CaFTa1	73.7	72	89.8	89.2																										
MtFTa2	65.9	65.3	77.4	78	80.2																									
PsFTa2	67.4	67.4	78.4	79	83	85.9																								
CaFTa2	69.1	68	78.4	77.8	84.1	81.9	86.9																							
GmFTa1/2a	72.6	70.9	84	83.4	87.4	76.1	79.4	80.6																						
GmFTa1/2b	72.6	70.3	81.7	81.1	84.6	76.7	78.9	80.6	94.3																					
MtFTa3	71.4	72.6	73.1	71.4	74.9	64.2	69.1	69.7	77.7	77.1																				
CaFTa3	71.2	68.9	70.8	68.5	73.6	64.2	66.9	68	75.7	75.7	86.4																			
GmFTa3a	72.6	72	73.3	73.9	77.8	70.1	73.3	73.9	80.6	78.9	85.1	84.3																		
GmFTa3b	72	71.4	73.3	72.7	77.3	68.4	72.2	71.6	79.4	78.9	83.4	83.7	90.9																	
GmFTa3c	72	71.4	73.7	73.1	77.7	68.8	72.6	72	79.4	78.9	83.4	83.6	91.4	100																
GmFTa3d	83.9	80.6	79.4	77.8	81	71.4	74.6	73	85.5	82.3	91.9	84.1	98.4	92.1	93.5															
MtFTc	61.7	61.7	58.9	58.9	59.4	54.5	57.1	57.1	60	59.4	64	61	66.3	61.1	61.1	64.5														
PsFTc	62.3	62.3	58.3	58.3	60.6	54.5	57.7	57.7	59.4	60.6	65.1	62.7	64.6	62.3	62.3	66.1	85.6													
CaFTc	63.3	63.8	59.3	58.8	62.7	55.6	58.8	59.3	62.7	62.1	67.2	65	66.7	65	65	67.7	83	87.5												
GmFTc1	64	64	57.7	59.4	60	56.3	59.4	59.4	61.7	59.4	65.7	61	65.7	60.6	60.6	66.1	74.1	74.1	77.3											
MtFTb1	68.2	67	69.5	68.9	70.1	63.5	65	63.8	68.2	69.3	70.5	68.2	68.9	69.5	69.9	75	61.9	64.2	65.7	64.2										
MtFTb2	64.8	65.9	67.2	66.7	67.8	64	65	63.8	68.8	70.5	69.9	65.9	68.4	67.8	68.2	73.4	59.1	61.4	63.5	61.4	88.8									
PsFTb1	61.9	64.8	66.1	65.5	66.1	61.2	61	59.9	64.2	65.3	65.9	63.7	63.3	65	65.3	64.1	57.4	60.2	61.8	60.8	86	86.5								
PsFTb2	60.8	64.8	66.1	66.1	66.7	62.9	62.7	61	65.3	66.5	66.5	64.2	66.1	66.7	67	67.2	58	60.8	62.4	60.8	86	89.3	93.8							
CaFTb	65	65	67.4	66.3	69.7	63.7	64	63.5	66.7	68.9	67.8	65.9	67.4	69.1	69.5	73.4	57.6	59.9	61.8	59.9	86.6	87.7	84.9	87.2						
GmFTb1	65.3	64.8	70.1	68.4	71.2	66.3	67.2	66.1	71	72.2	71	67	71.2	70.1	70.5	71.9	60.2	61.9	62.9	60.6	79.7	80.9	77	78.1	78.2					
GmFTb2	25	25.6	25.4	25.4	26.6	23.6	25.4	24.3	26.7	26.7	26.7	25.1	26.6	27.1	27.3	50.8	21	21.6	22.5	21.4	29.9	27.9	28.5	27.9	28.3	33.7				
GmFTb3	61.6	66.7	59.9	61.6	62.1	56.2	60.5	61.6	63.8	66.1	67.2	63.1	67.8	66.1	66.1	71.4	59.3	59.9	62.6	61.9	68.8	70.1	67.8	67.2	65.7	70.5	25.4			
GmFTb4	61.8	68	61.2	61.2	62.9	57	60.7	60.7	65.7	65.7	66.9	62.8	65.2	64.6	64.6	70.3	56.7	55.1	57.2	58.9	68.8	68.5	66.9	66.9	67.6	71.8	26.4	74.7		
GmFTb5	63.1	63.1	64.2	63.6	63.6	58.2	62.5	61.4	62.5	64.8	65.9	63.5	68.2	68.8	68.8	72.6	57.4	57.4	58.4	55.4	69.1	71.6	68.8	69.3	68.9	70.3	27.8	68.8	63.8	
GmFT-like	44.8	45.5	44.8	44.2	45.5	41.2	42.4	44.2	41.8	42.4	40	40	41.8	41.2	41.2	50.9	37.6	38.2	39.4	35.8	39.8	38	37.3	37.3	39.2	39.8	15.8	41.6	39.8	41

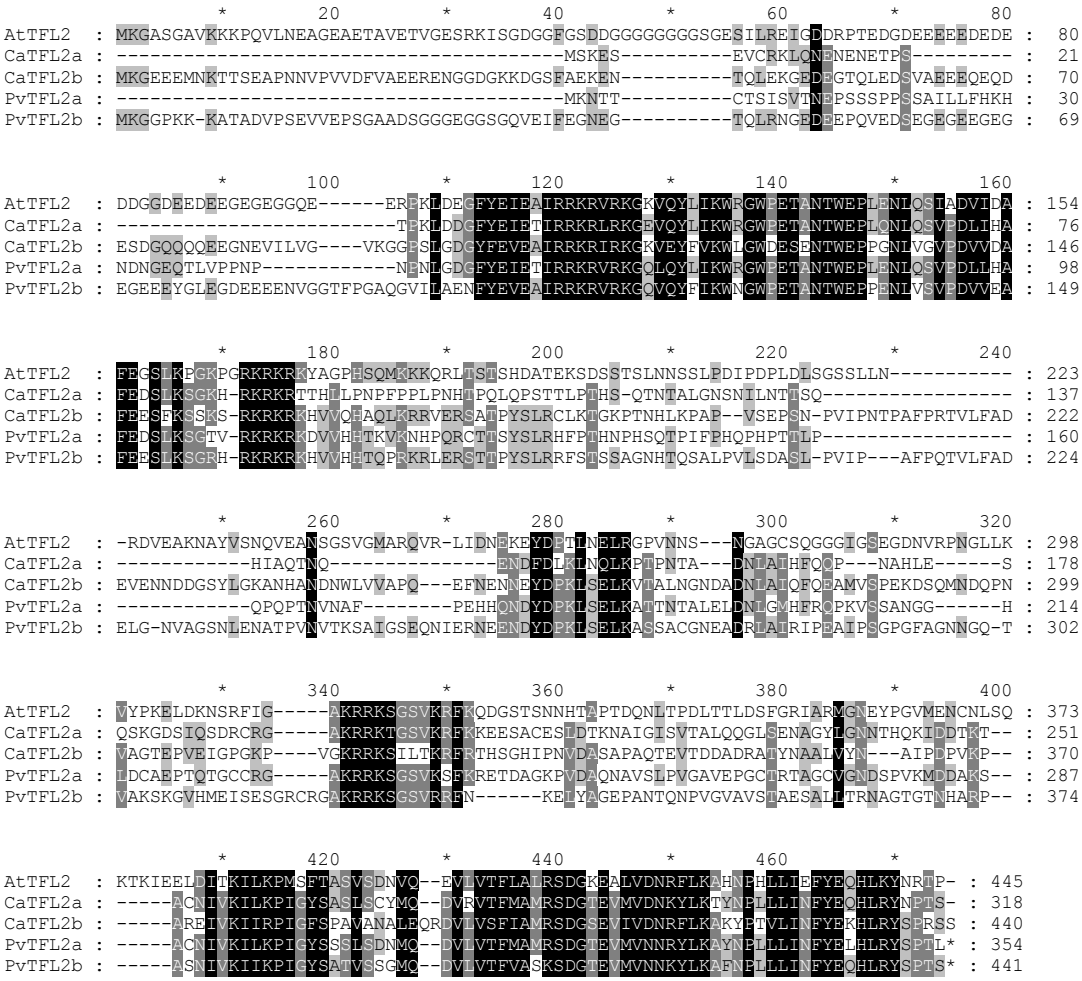
Appendix 3.26 Homology matrix (%) derived from the multiple sequence alignment of *FLOWERING PROMOTING FACTOR 1* (FPF1) and FPF1-like genes in the species *Arabidopsis thaliana* (At) and *Cicer arietinum* (Ca). Protein sequences from accessions listed in the table were aligned using MAFFT v1.3.3 in Geneious 8 software.

	Accession	AtFPF1	AtFPF1-like	AtFPF1-like	CaFPF1a	CaFPF1b	CaFPF1c	CaFPF1d
AtFPF1	AT5G24860		79.6	66.1	70.0	58.4	61.9	42.0
AtFPF1-like1	AT5G10625	79.6		64.5	67.9	58.9	60.7	44.3
AtFPF1-like2	AT4G31380	66.1	64.5		62.1	56.5	55.6	42.4
CaFPF1a	LOC101489303	70.0	67.9	62.1		58.9	61.1	42.7
CaFPF1b	LOC101491798	58.4	58.9	56.5	58.9		73.0	40.2
CaFPF1c	LOC101494723	61.9	60.7	55.6	61.1	73.0		46.6
CaFPF1d	LOC101507122	42.0	44.3	42.4	42.7	40.2	46.6	



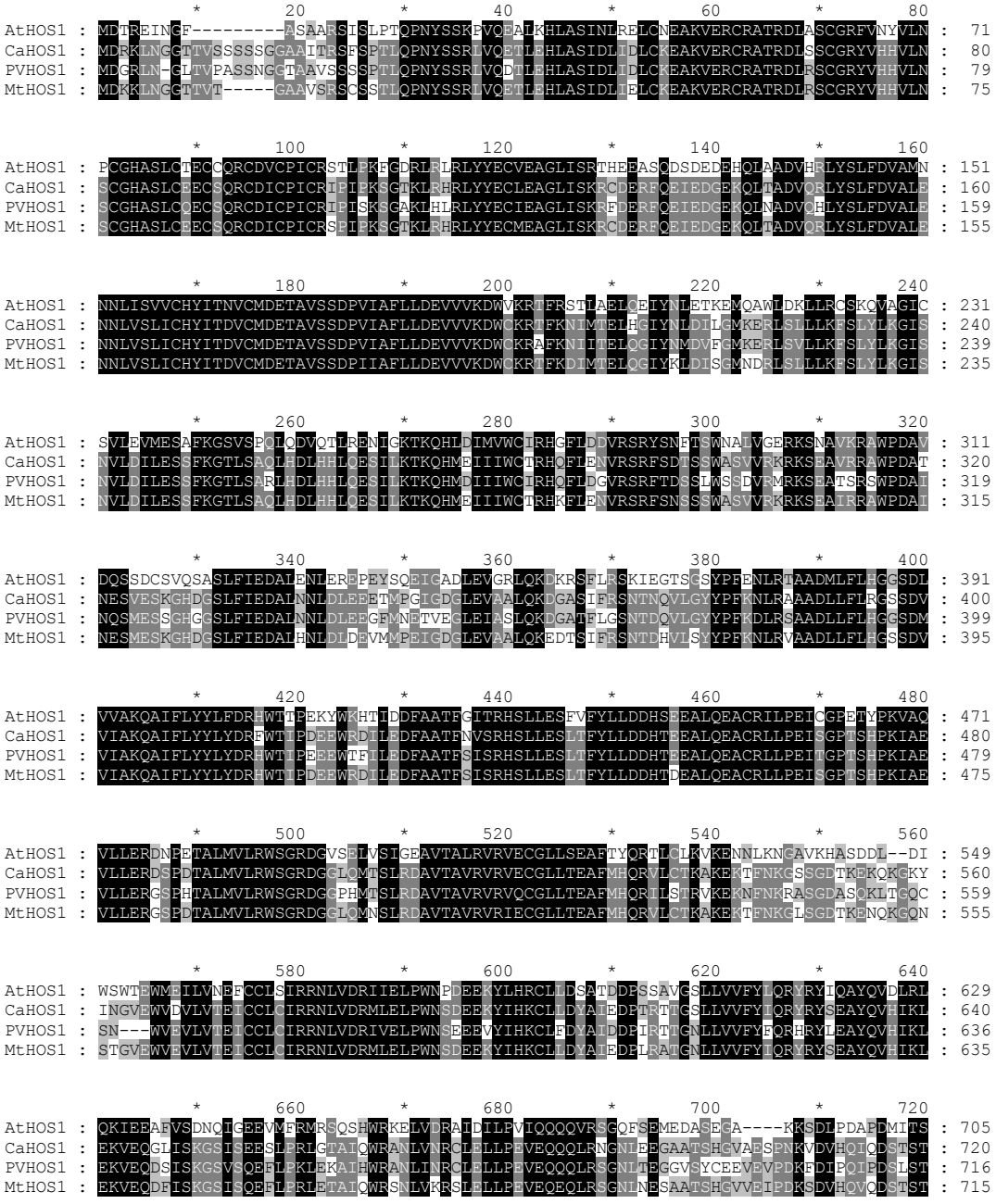
Appendix 3.27 Homology matrix (%) derived from the multiple sequence alignment of *TERMINAL FLOWER 2* (TFL2) in the species *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca) and *Phaseolus vulgaris* (Pv). Protein sequences from accessions listed in the table were aligned using ClustalW in Geneious 8 software.

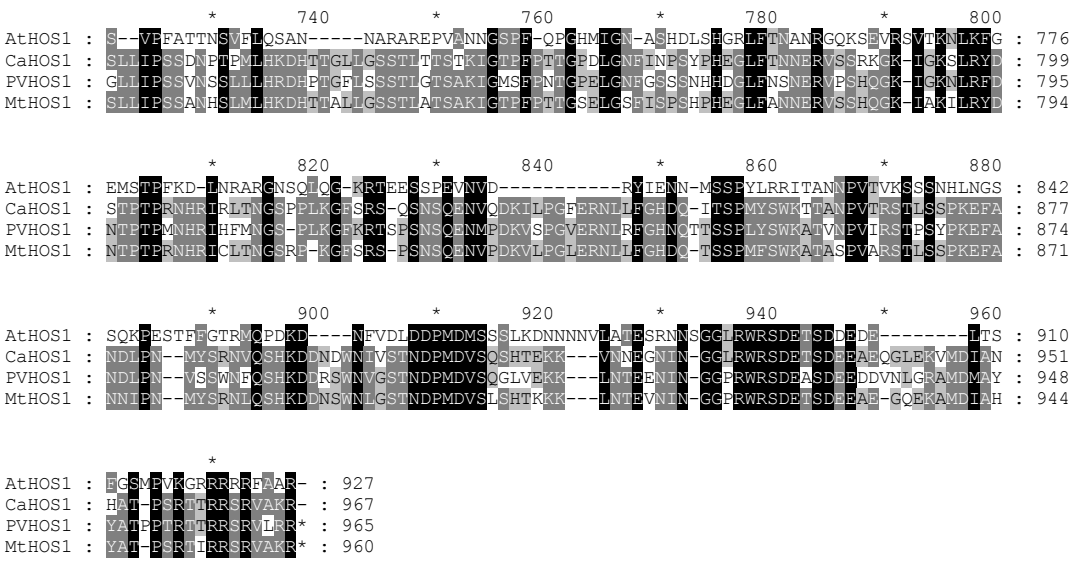
	Accession	AtTFL2	CaTFL2a	CaTFL2b	PvTFL2a	PvTFL2b
AtTFL2	AT5G17690		31.4	28.0	33.1	31.4
CaTFL2a	LOC101503633	31.4		29.4	51.1	33.5
CaTFL2b	LOC101501646	28.0	29.4		35.0	44.5
PvTFL2a	Phvul.009G117500	33.1	51.1	35.0		39.0
PvTFL2b	Phvul.010G034600	31.4	33.5	44.5	39.0	



Appendix 3.28 Identity matrix (%) derived from the multiple sequence alignment of *HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1* (HOS1) genes in the species *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca), *Phaseolus vulgaris* (Pv) and *Medicago truncatula* (Mt). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

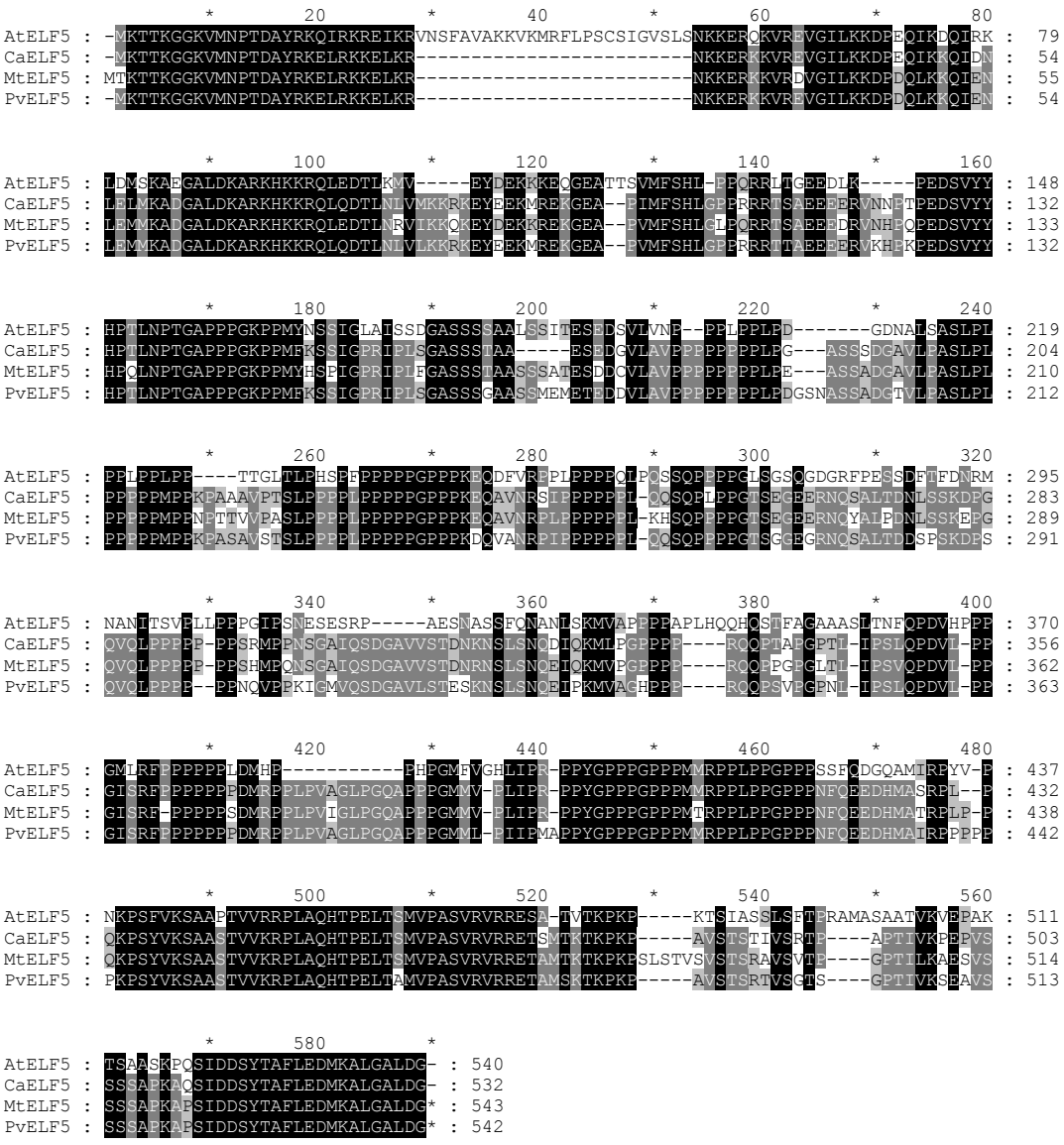
	Accession	AtHOS1	CaHOS1	PVHOS1	MtHOS1
AtHOS1	AT2G39810		49.3	48.6	49.4
CaHOS1	LOC101509472	49.3		77.1	87.2
PVHOS1	Phvul.002G056400	48.6	77.1		77.0
MtHOS1	Medtr5g012940	49.4	87.2	77.0	




AtHOS1 : S--VPFATTNSVPLQSAN-----NARAREPVPNNGSPP-QPGHMIGN-ASHDLSHGRLFTNANRGQKSEVRSVTKNLKFG : 776
CaHOS1 : SLLIPSSDNPTPMLHKDHTTGLIGSSTLTTSTKIGTPPPTTGPDLGNPINPSYPHEGLFTNNERVSSRKGK-IGKSLRYD : 799
PVHOS1 : GLLIPSSVNSSLILHRDHPTGFISSSTIGTSAKIGMSPPNTGPELGNPGSSSNHHDGLFNSNERVPSHQGK-IGKNLRFD : 795
MtHOS1 : SLLIPSSANHSLMLHKDHTTALIGSSTLATSAKIGTPPPTTGSELGSPISPSHPHEGLFANNERVSSHQGK-IAKILRYD : 794
AtHOS1 : EMSTPFKD-LNRARGNSQLQG-KRTEPSSPEVNVD----------RYIENN-MSSPYLRRITANNPVTVKSSSNHLNGS : 842
CaHOS1 : STPTPRNHRIRLTNGSPPLKGPSRS-QSNSQENVQDKILPGPERNLLPGHDQ-ITSPMYSWKTTANPVTRSTLSSPKEFA : 877
PVHOS1 : NTPTPMNHRIHFMNGS-PLKGPKRTSPSNSQENMPDKVSPGVERNLRPGHNQTTSSPLYSWKATVNPVIRSTPSYPKEFA : 874
MtHOS1 : NTPTPRNHRICLTNGSRP-KGPSRS-PSNSQENVPDKVLPGLERNLLPGHDQ-TSSPMFSWKATASPVARSTLSSPKEFA : 871
AtHOS1 : SQKPESTPPGTRMQPDKD----NFVDLDDPMDMSSSLKDNNNNVLATPSRNNSGGLRWRSDETSDDEDE--------LTS : 910
CaHOS1 : NDLPN--MYSRNVQSHKDDNDWNIVSTNDPMDVSQSHTEKK---VNNPGNIN-GGLRWRSDEISDEEAEQGLEKVMDIAN : 951
PVHOS1 : NDLPN--VSSWNPQSHKDDRSWNVGSTNDPMDVSQGLVEKK---LNTPENIN-GGPRWRSDEASDEEDDVNLGRAMDMAY : 948
MtHOS1 : NNIPN--MYSRNLQSHKDDNSWNLGSTNDPMDVSLSHTKKK---LNTPVNIN-GGPRWRSDEISDEEAE-GQEKAMDIAH : 944
AtHOS1 : PGSMPVKGRRRRRFAAR- : 927
CaHOS1 : HAT-PSRTTRRSRVAKR- : 967
PVHOS1 : YATPPTRTTRRSRVLRR* : 965
MtHOS1 : YAT-PSRTIRRSRVAKR* : 960

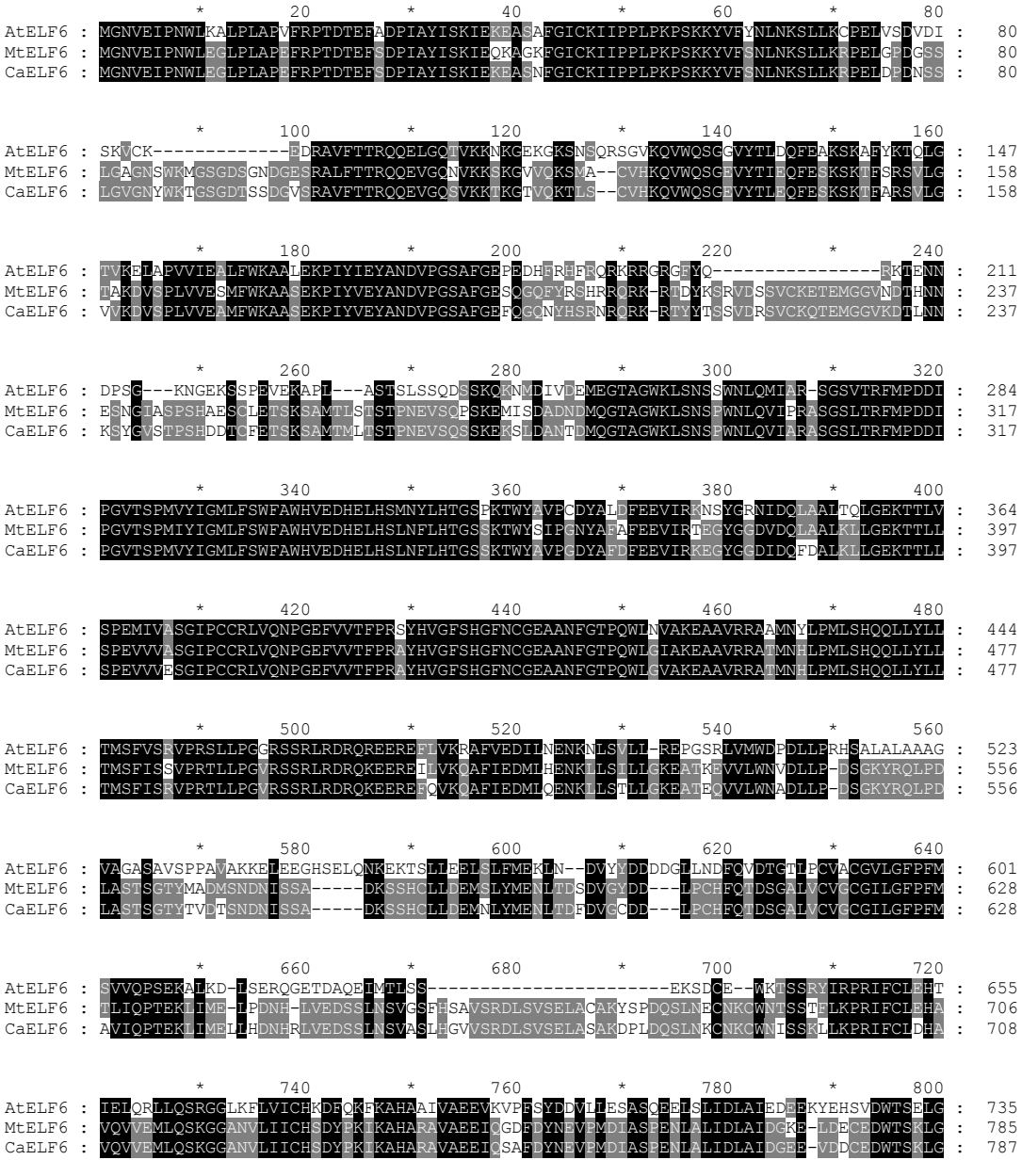
Appendix 3.29 Identity matrix (%) derived from the multiple sequence alignment of *EARLY FLOWERING 5* (ELF5) gene in the species *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca), *Medicago truncatula* (Mt) and *Phaseolus vulgaris* (Pv). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

	Accession	AtELF5	CaELF5	MtELF5	PvELF5
AtELF5	AT5G62640		55.0	54.4	54.2
CaELF5	LOC101497894	55.0		86.2	86.0
MtELF5	Medtr4g107410	54.4	86.2		82.0
PvELF5	Phvul.003G194500	54.2	86.0	82.0	



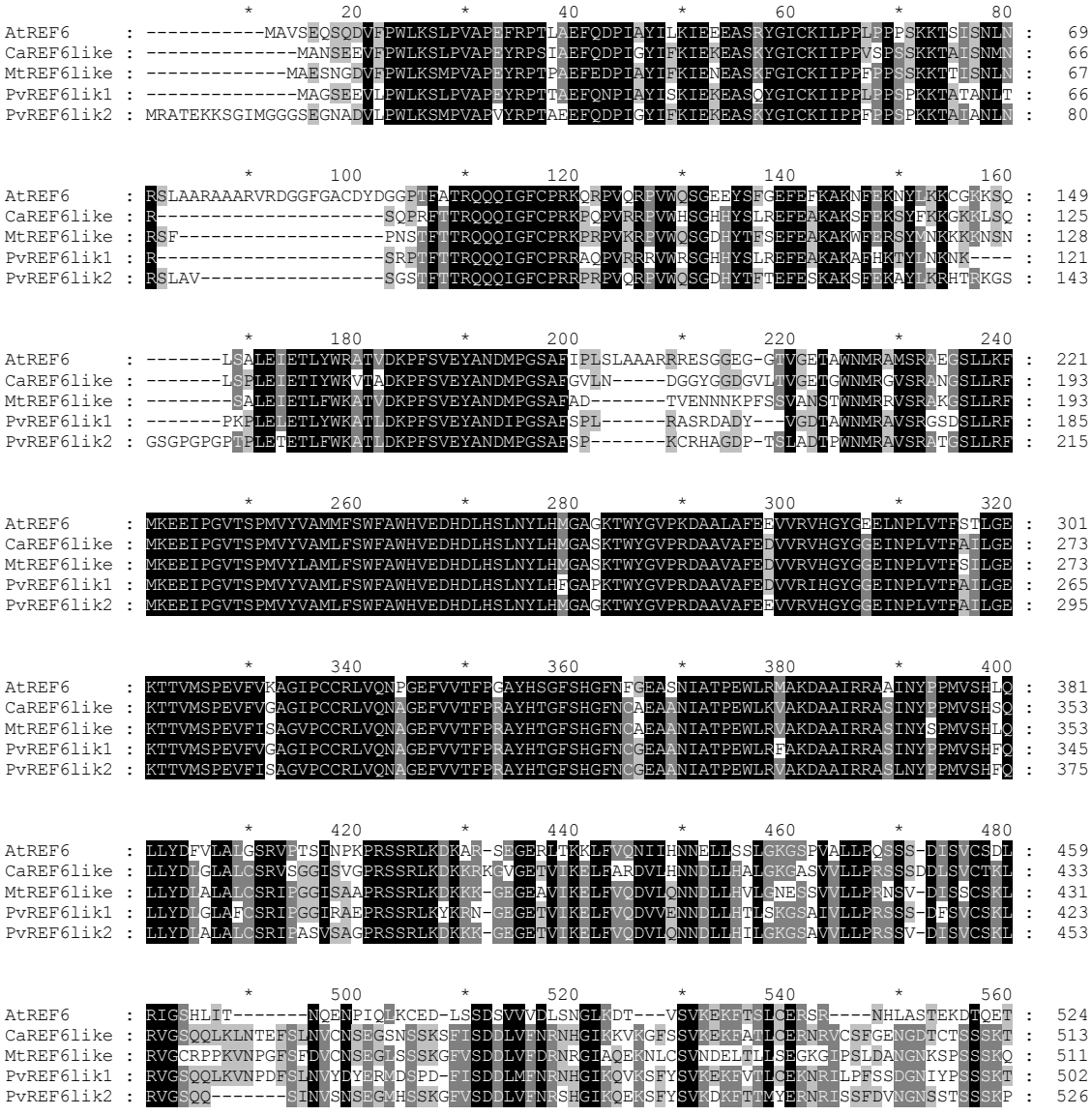
Appendix 3.30 Identity matrix (%) derived from the multiple sequence alignment of *EARLY FLOWERING 6* (ELF6) genes in the species *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt) and *Cicer arietinum* (Ca). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

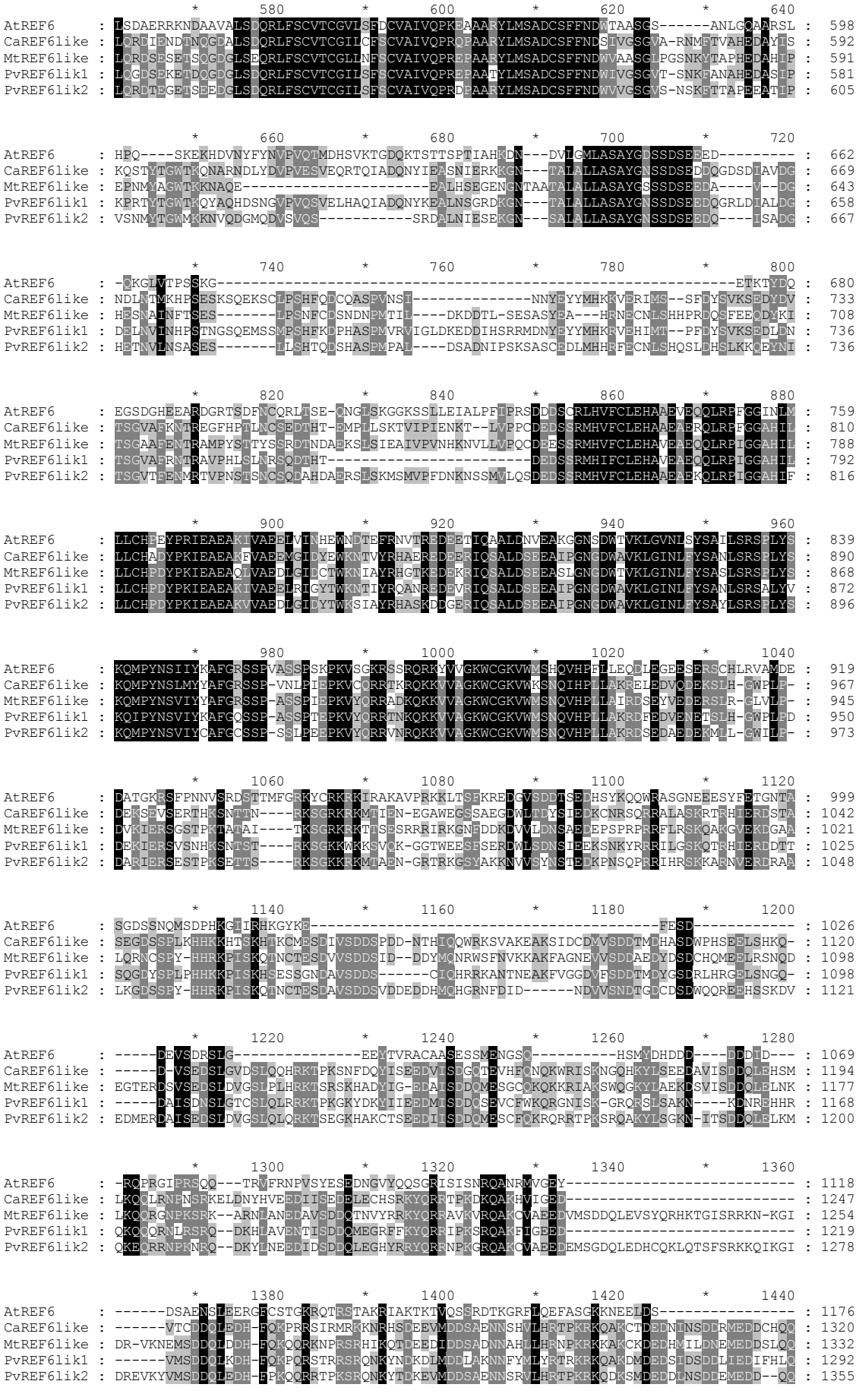
	Accession	AtELF6	MtELF6	CaELF6
AtELF6	AT5G04240		42.1	45.2
MtELF6	Medtr1g094740	42.1		77.6
CaELF6	LOC101509509	45.2	77.6	

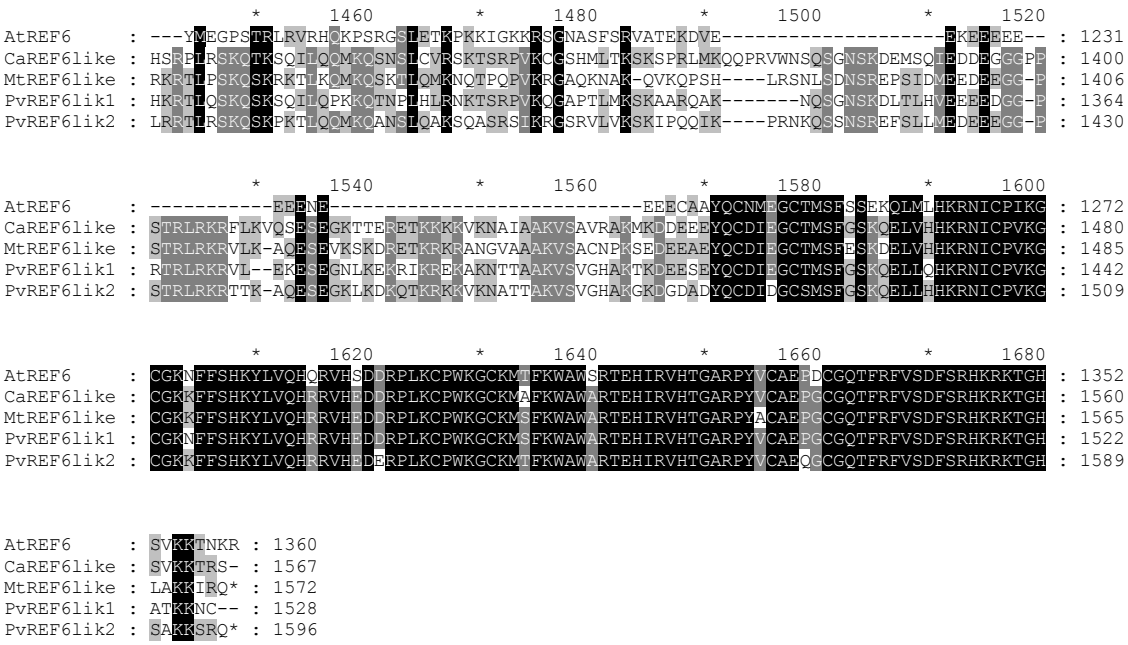


Appendix 3.31 Identity matrix (%) derived from the multiple sequence alignment of RELATIVE OF EARLY FLOWERING 6 (REF6) genes in the species *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca), *Medicago truncatula* (Mt) and *Phaseolus vulgaris* (Pv). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

	Accession	AtREF6	CaREF6-like	REF6-like	PvREF6-like1	PvREF6-like2
AtREF6	AT3G48430		44.1	42.3	43.3	42.7
CaREF6-like	LOC101495646	44.1		60.9	68.8	64.1
MtREF6-like	Medtr3g075210	42.3	60.9		59.3	68.9
PvREF6-like1	Phvul.003G207700	43.3	68.8	59.3		63.1
PvREF6-like2	Phvul.009G169700	42.7	64.1	68.9	63.1	




AtREF6
CaREF6like
MtREF6like
PvREF6lik1
PvREF6lik2


* 1460 * 1480 * 1500 * 1520
AtREF6 : ---YMEGPSTRLRVRHCKPSRGSIETKPKKIGKKRSGNASFSRVATEKDVE----------EKEEEEE-- : 1231
CaREF6like : HSFPLRSKQTKSQILQQMKQSNSLCVRSKTSRPVKCGSHMLTKSKSPRLMKQQPRVWNSQSGNSKDEMSQIEDDEGGGPP : 1400
MtREF6like : RKRTLPSKQSKRKTLKQMKQSKTLQMKNQTPQPVKRGAQKNAK-QVKQPSH----LRSNLSDNSREPSIDMEEDEEGG-P : 1406
PvREF6lik1 : HKRTLQSKQSKSQILQPKKQTNPTHLRNKTSRPVKQGAPTLMKSKAARQAK-------NQSGNSKDLTLHVEEEEDGG-P : 1364
PvREF6lik2 : LRRTLRSKQSKPKTLQQMKQANSLQAKSQASRSTKRGSRVLVKSKIPQQIK----PRNKQSSNSREFSLLMEDEEEGG-P : 1430

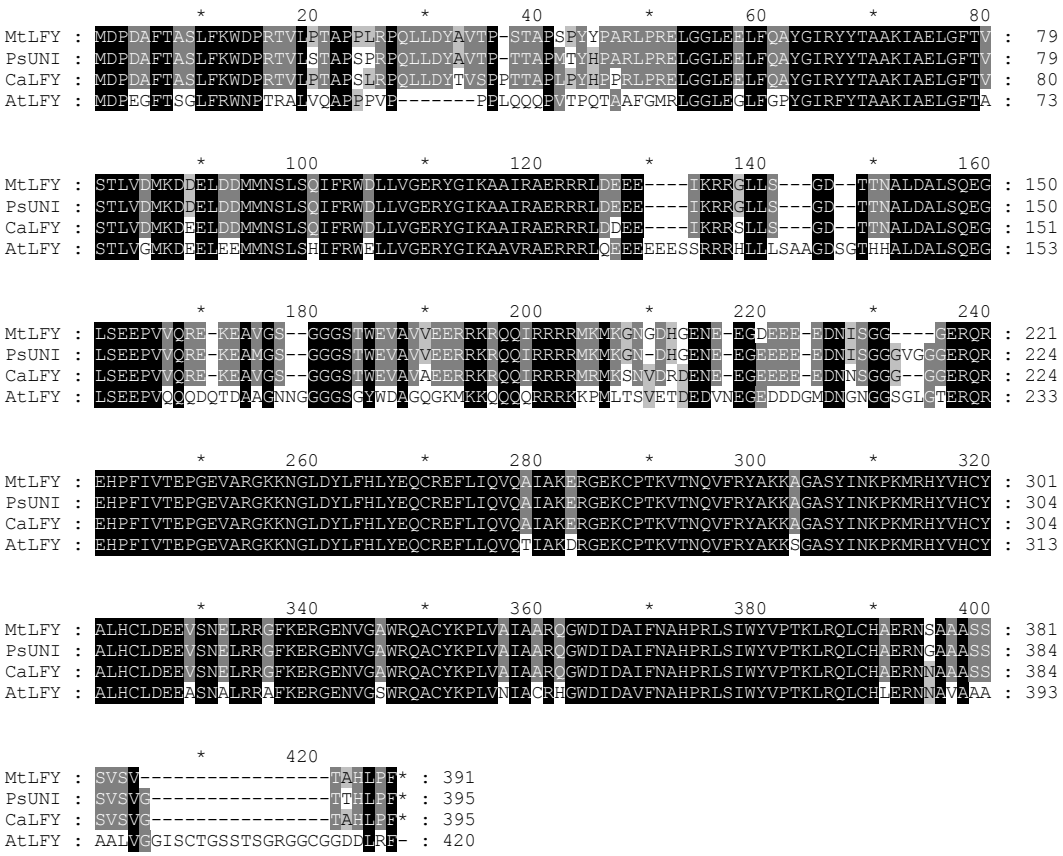
* 1540 * 1560 * 1580 * 1600
AtREF6 : -----------EEENE----------EEECAAYQCNMEGCTMSFSSEKQLMLHKRNICPIKG : 1272
CaREF6like : STRLRKRFLKVQSESEGKTTERETKKKKVKNAIAAKVSAVRAKMKDDEEEYQCDIEGCTMSFGSKQELVHHKRNICPVKG : 1480
MtREF6like : STRLRKRVLK-AQESEVKSKDRETKRKRANGVAAAKVSACNPKSEDEEAEYQCDIEGCTMSFESKDELVHHKRNICPVKG : 1485
PvREF6lik1 : RTRLRKRVL--EKESEGNLKEKRIKREKAKNTTAAKVSVGHAKTKDEESEYQCDIEGCTMSFGSKQELLQHKRNICPVKG : 1442
PvREF6lik2 : STRLRKRTTK-AQESEGKLKDKQTKRKKVKNATTAKVSVGHAKGKDGDADYQCDIDGCSMSFGSKQELLHHKRNICPVKG : 1509

* 1620 * 1640 * 1660 * 1680
AtREF6 : CGKNFFSHKYLVQHQRVHSDDRPLKCPWKGCKMTFKWAWSRTEHIRVHTGARPYVCAEIDCGQTFRFVSDFSRHKRKTGH : 1352
CaREF6like : CGKKFFSHKYLVQHRRVHEDDRPLKCPWKGCKMAFKWAWARTEHIRVHTGARPYVCAEPGCGQTFRFVSDFSRHKRKTGH : 1560
MtREF6like : CGKKFFSHKYLVQHRRVHEDDRPLKCPWKGCKMSFKWAWARTEHIRVHTGARPYACAEPGCGQTFRFVSDFSRHKRKTGH : 1565
PvREF6lik1 : CGKNFFSHKYLVQHRRVHEDDRPLKCPWKGCKMSFKWAWARTEHIRVHTGARPYVCAEPGCGQTFRFVSDFSRHKRKTGH : 1522
PvREF6lik2 : CGKKFFSHKYLVQHRRVHEDERPLKCPWKGCKMTFKWAWARTEHIRVHTGARPYVCAEQGCGQTFRFVSDFSRHKRKTGH : 1589

AtREF6 : SVKKTNKR : 1360
CaREF6like : SVKKTRS- : 1567
MtREF6like : LAKKIRQ* : 1572
PvREF6lik1 : ATKKNC-- : 1528
PvREF6lik2 : SAKKSRQ* : 1596

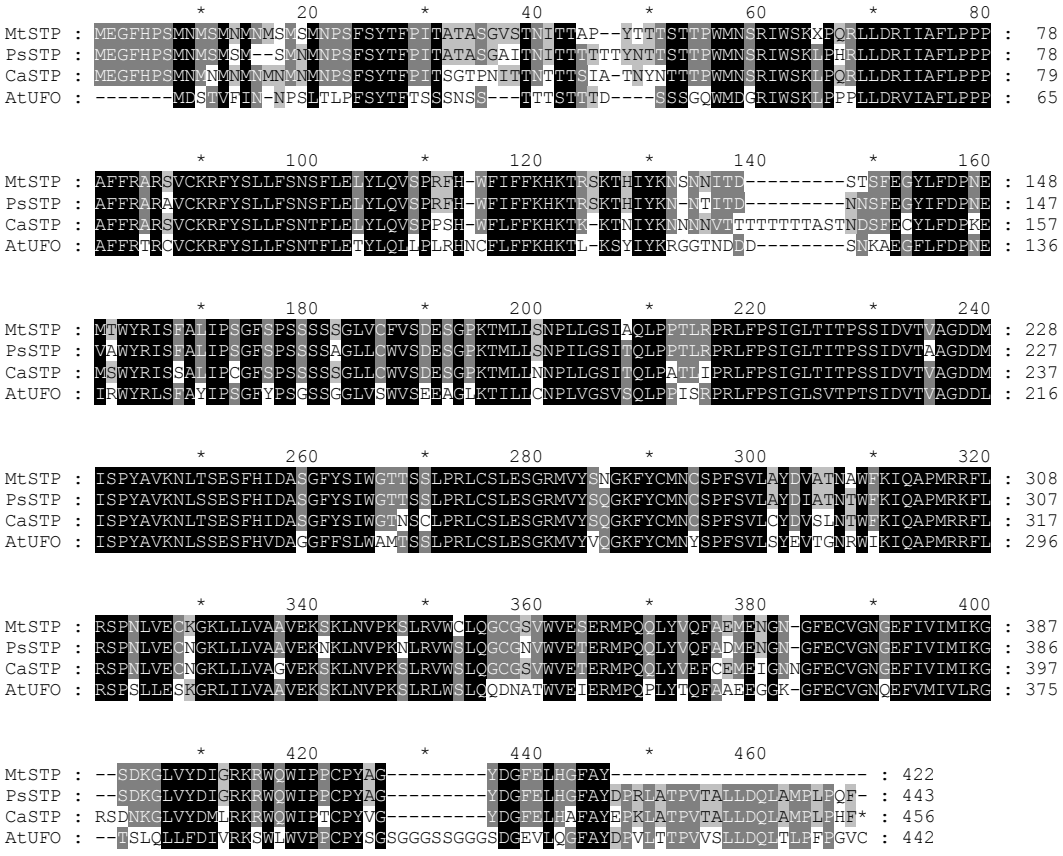
Appendix 3.32 Identity matrix (%) derived from the multiple sequence alignment of *LEAFY* (LFY) genes in *Cicer arietinum* (Ca), *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt) and *Pisum sativum* (Pv). Protein sequences from accessions listed in the table were aligned using ClustalW in Geneious 8 software.

	Accession	MtLFY	PsUNI	CaLFY	AtLFY
MtLFY	Medtr3g098560		95.7	94.1	64.4
PsUNI	AF010190	95.7		94.2	65.3
CaLFY	LOC101493650	94.1	94.2		65.3
AtLFY	AT5G61850	64.4	65.3	65.3	



Appendix 3.33 Identity matrix (%) derived from the multiple sequence alignment of *UNUSUAL FLORAL ORGANS* (UFO)/*Stamina pistilloidia* (STP) genes in *Cicer arietinum* (Ca), *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt) and *Pisum sativum* (Pv). Protein sequences from accessions listed in the table were aligned using ClustalW in Geneious 8 software.

	Accession	MtSTP	PsSTP	CaSTP	AtUFO
MtSTP	Medtr4g094748		90.8	82.5	63.4
PsSTP	AF004843	90.8		81.2	62.5
CaSTP	LOC101505902	82.5	81.2		58.9
AtUFO	AT1G30950	63.4	62.5	58.9	



Appendix 3.34 Maximum parsimony tree derived from the alignment of the flower identity genes *APETALA2* (AP2) and the AP2-like genes *TARGET OF EAT* (TOE), *SCHNARCHZAPFEN* (SNZ) and *SCHLAFMUTZE* (SMZ) in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Pisum sativum* (Ps), *Glycine max* (Gm) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using MAFFT and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications.

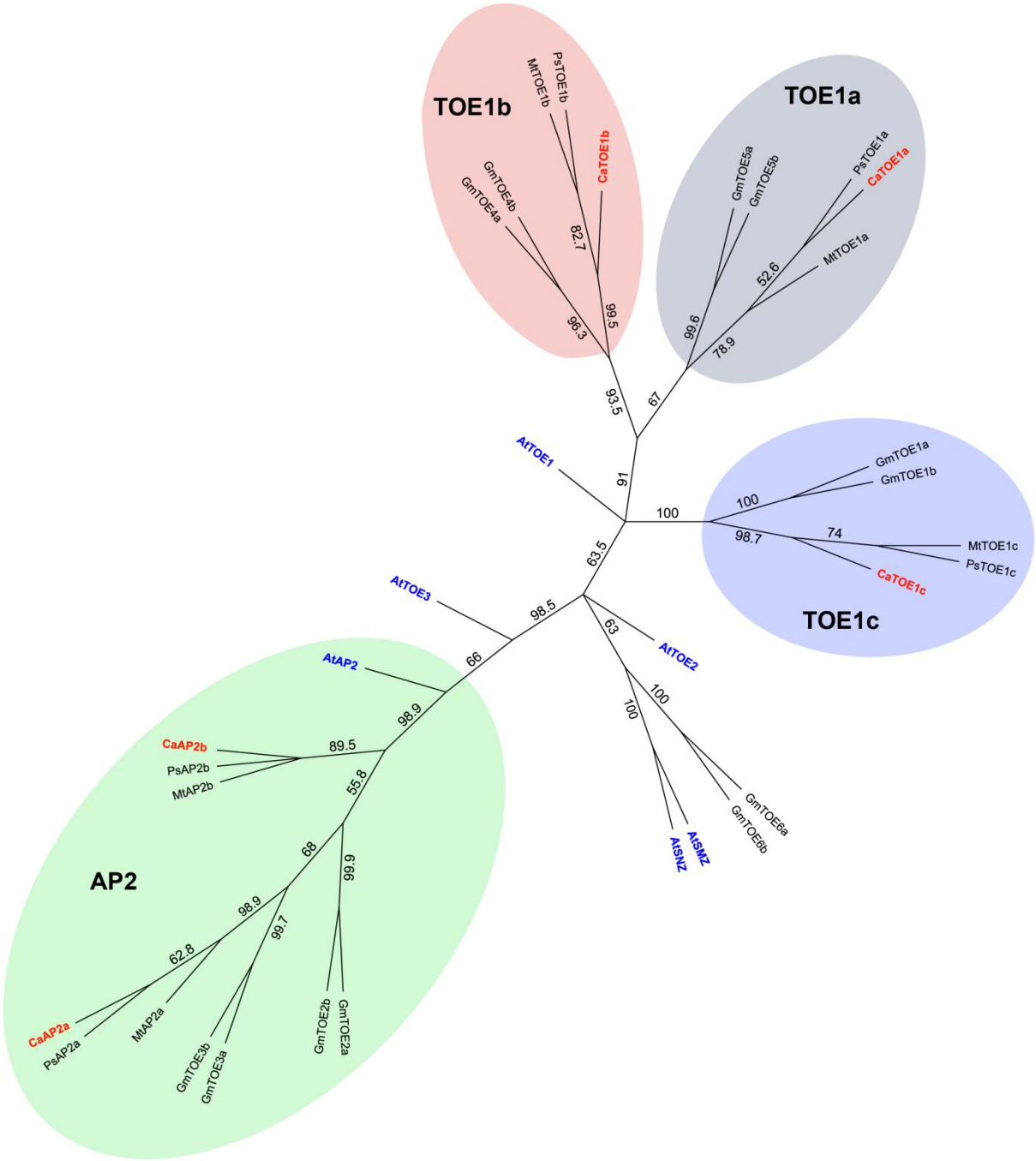


Table 1 Accession number of the protein sequences used in the alignment and tree construction of five genes belonging to the AP2-family.

Arabidopsis thaliana	
AP2	AT4G36920
TOE1	AT2G28550
TOE2	AT5G60120
TOE3	AT5G67180
SNZ	AT2G39250
SMZ	AT3G54990

Glycine max	
TOE1a	Glyma.03G177500
TOE1b	Glyma.19G178200
TOE2a	Glyma.17G170300
TOE2b	Glyma.05G091200
TOE3a	Glyma.11G053800
TOE3b	Glyma.01G188400
TOE4a	Glyma.15G044400
TOE4b	Glyma.13G329700
TOE5a	Glyma.12G073300
TOE5b	Glyma.U022100
TOE6b	Glyma.02G087400
TOE6a	Glyma.10G116600

Medicago truncatula	
AP2a	Medtr5g016810
AP2b	Medtr4g094868
TOE1c	Medtr7g100590
TOE1b	Medtr2g093060
TOE1a	Medtr4g061200

Pisum sativum	
AP2a	PsCam050516
AP2b	PsCam020804
TOE1c	PsCam006792
TOE1b	PsCam000232
TOE1a	PsCam042790

Cicer arietinum	
AP2a	LOC101512375
AP2b	LOC101495928
TOE1a	LOC101489539
TOE1b	LOC101497653
TOE1c	LOC101511414

Appendix 3.35 Maximum parsimony tree derived from the alignment of HOMOLOGUE OF TRITHORAX (ATX) genes in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt), *Phaseolus vulgaris* (Pv) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using MAFFT and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications.

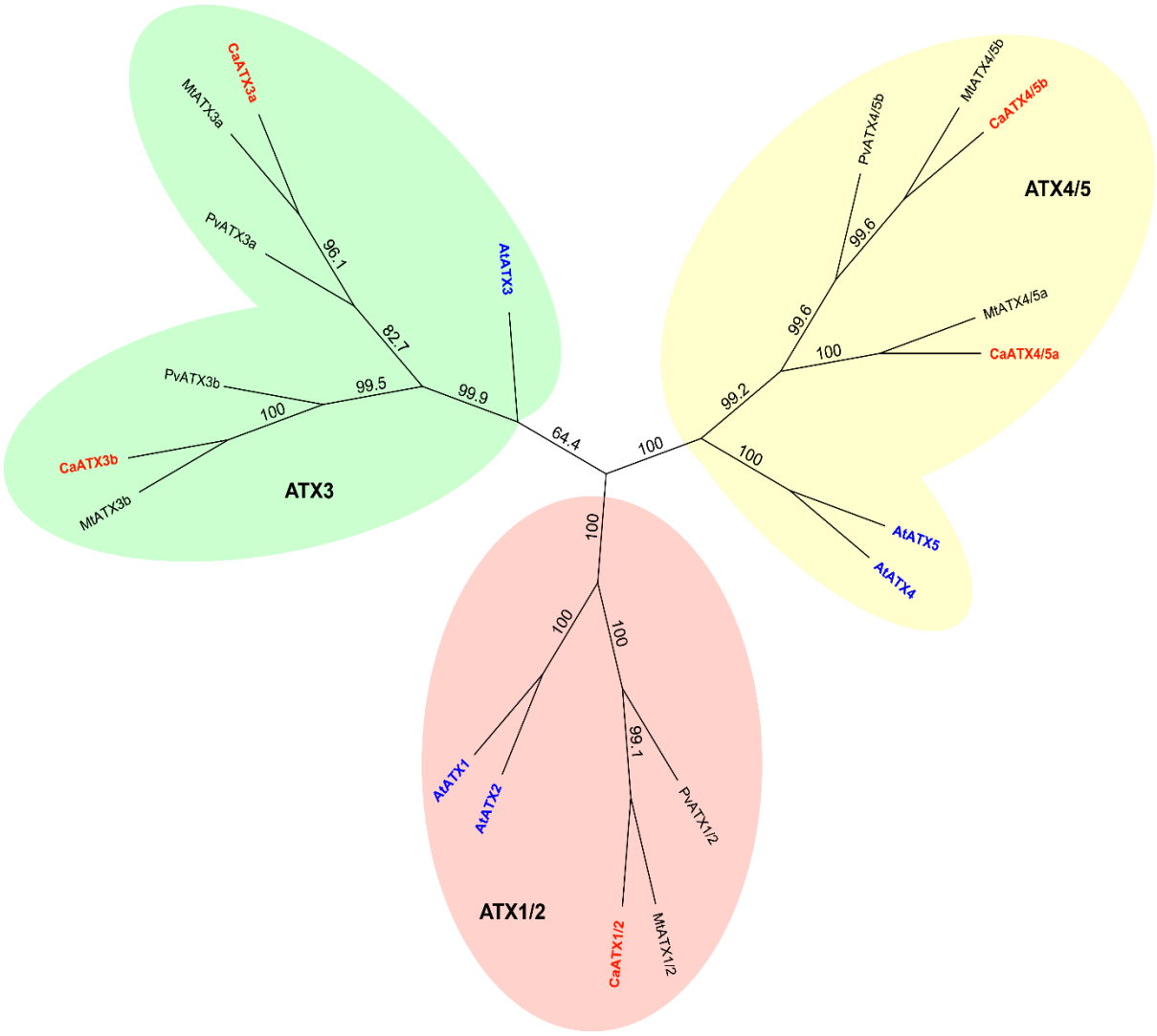


Table 1 Accession number of the protein sequences used in the alignment and tree construction of ATX genes.

Arabidopsis thaliana	
AtATX1	AT2G31650
AtATX2	AT1G05830
AtATX3	AT3G61740
AtATX4	AT4G27910
AtATX5	AT5G53430

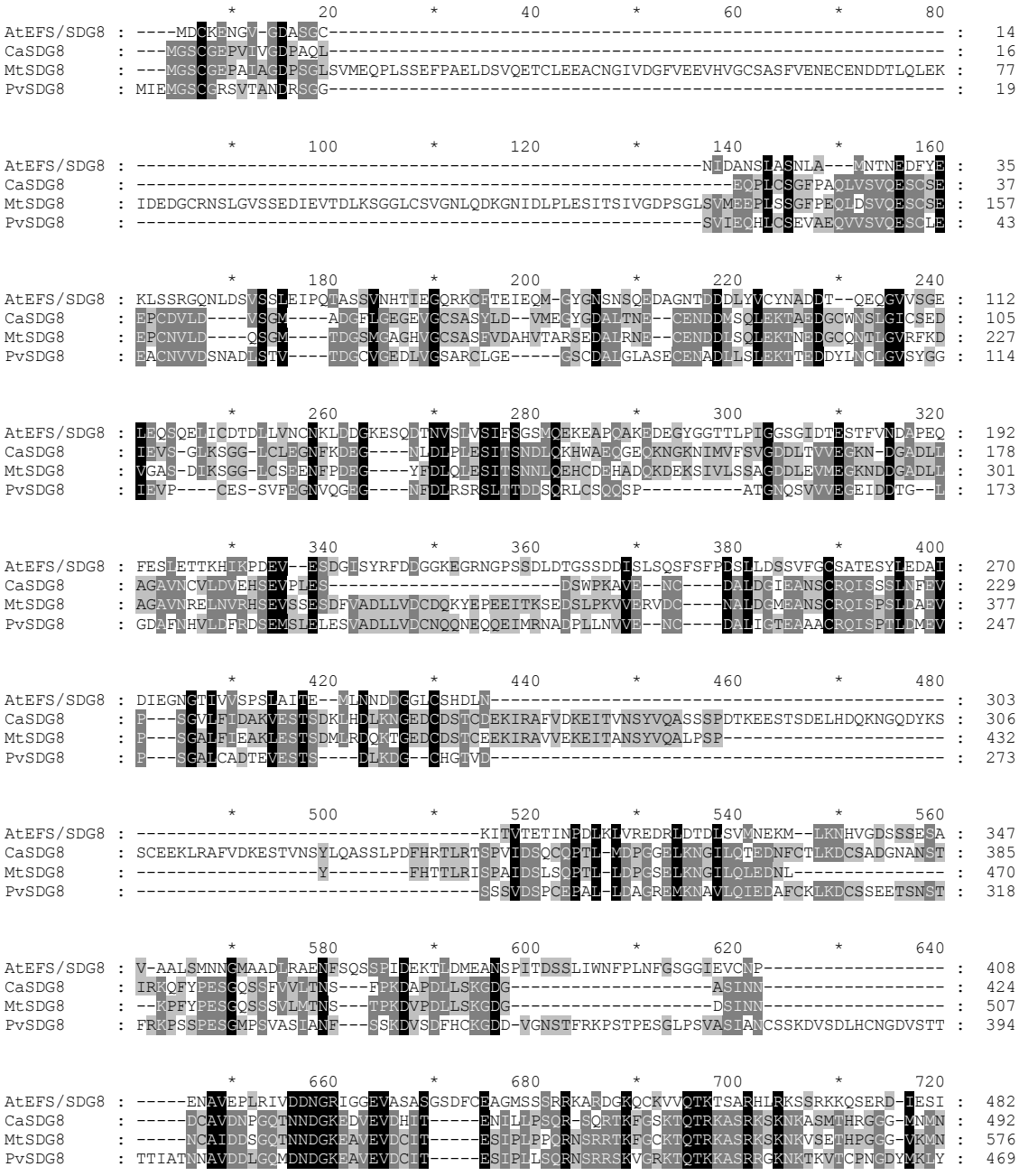
Cicer arietinum	
CaATX1/2	LOC101490018
CaATX3a	LOC101496041
CaATX3b	LOC101508120
CaATX4/5a	LOC101499063
CaATX4/5b	LOC101507097

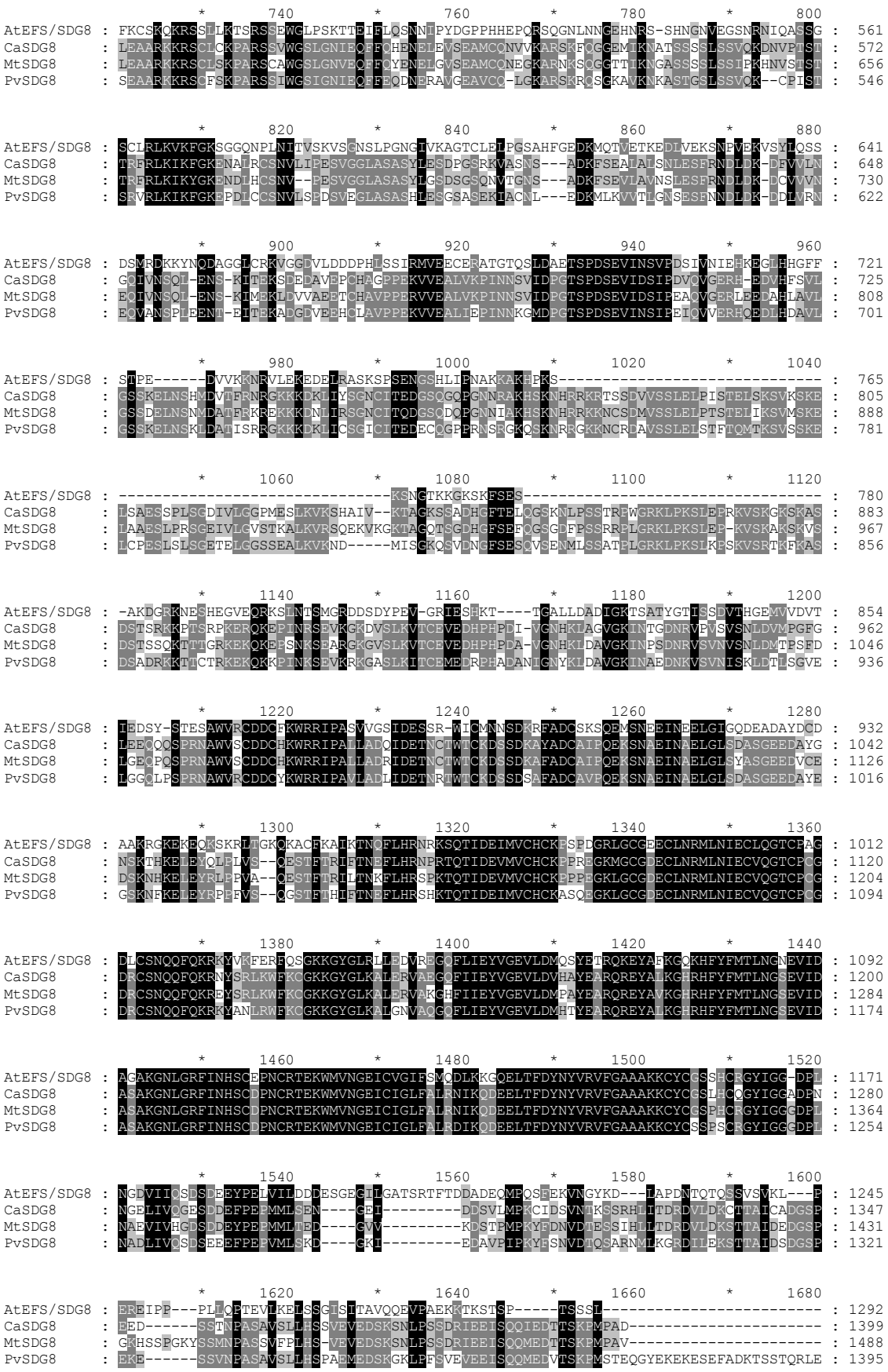
Medicago truncatula	
MtATX1/2	Medtr7g021365
MtATX3a	Medtr8g027725
MtATX3b	Medtr7g117355
MtATX4/5a	Medtr1g008230
MtATX4/5b	Medtr3g091310

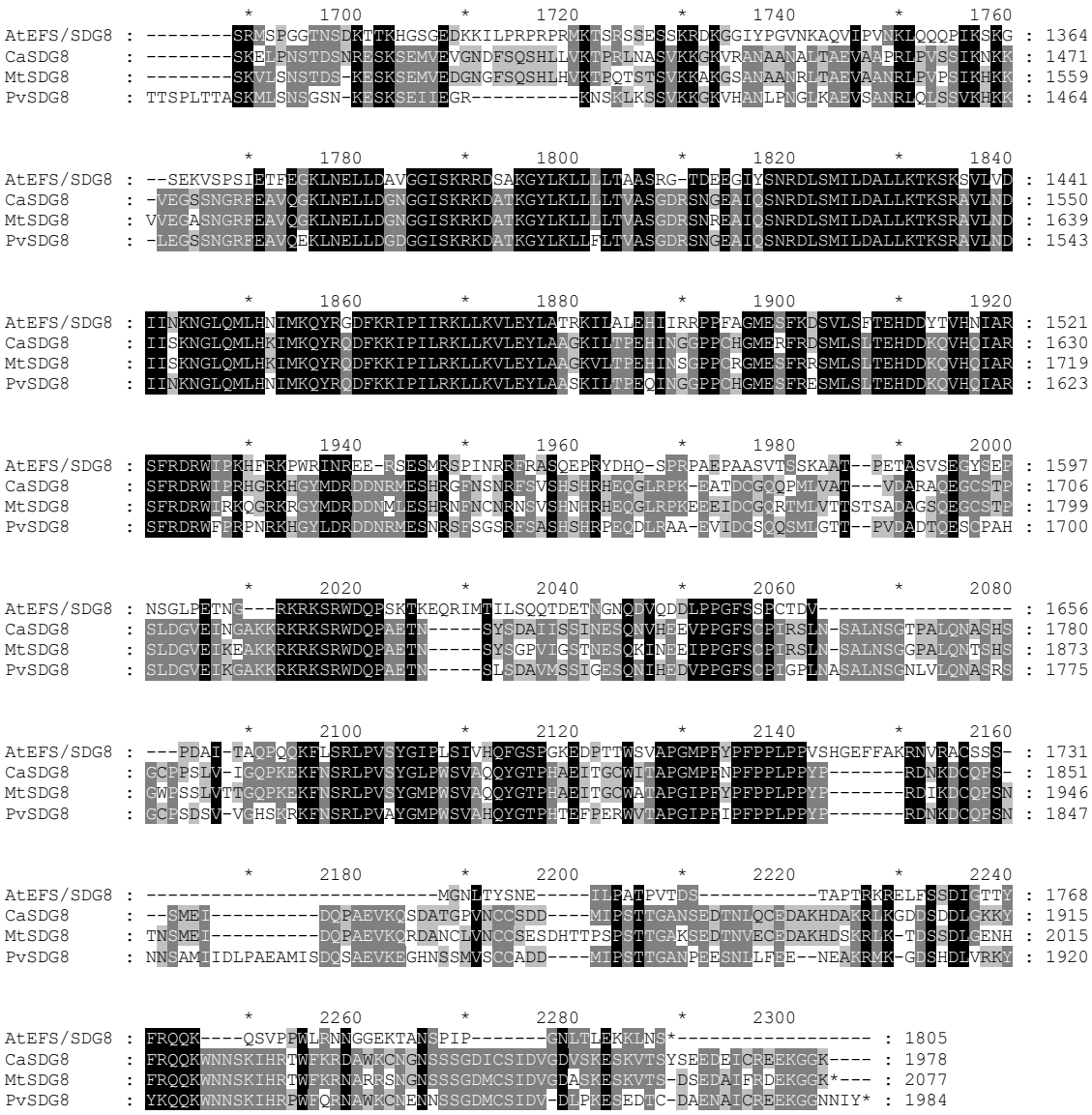
Phaseolus vulgaris	
PvATX1/2	Phvul.008G018500
PvATX3a	Phvul.010G113900
PvATX3b	Phvul.001G208500
PvATX4/5b	Phvul.009G053400

Appendix 3.36 Identity matrix (%) derived from the multiple sequence alignment of *EARLY FLOWERING IN SHORT DAYS* (EFS) genes in four species: *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca), *Medicago truncatula* (Mt) and *Phaseolus vulgaris* (Pv). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

	Accession	AtEFS	CaSDG8	MtSDG8	PvSDG8
AtEFS/SDG8	AT1G77300		29.7	28.4	29.9
CaSDG8	LOC101503412	29.7		71.0	60.0
MtSDG8	Medtr3g088625	28.4	71.0		56.9
PvSDG8	Phvul.009G130100	29.9	60.0	56.9	

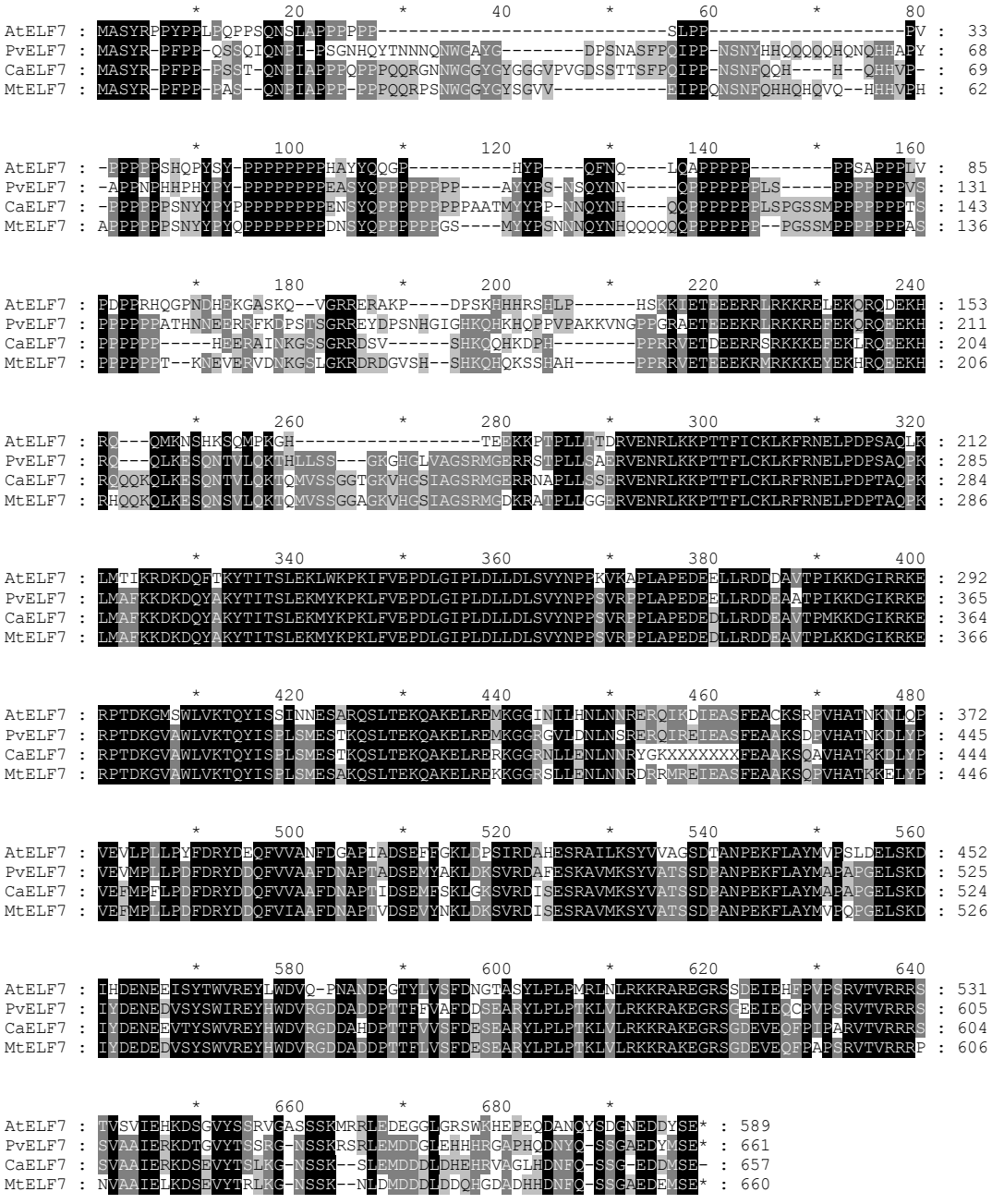






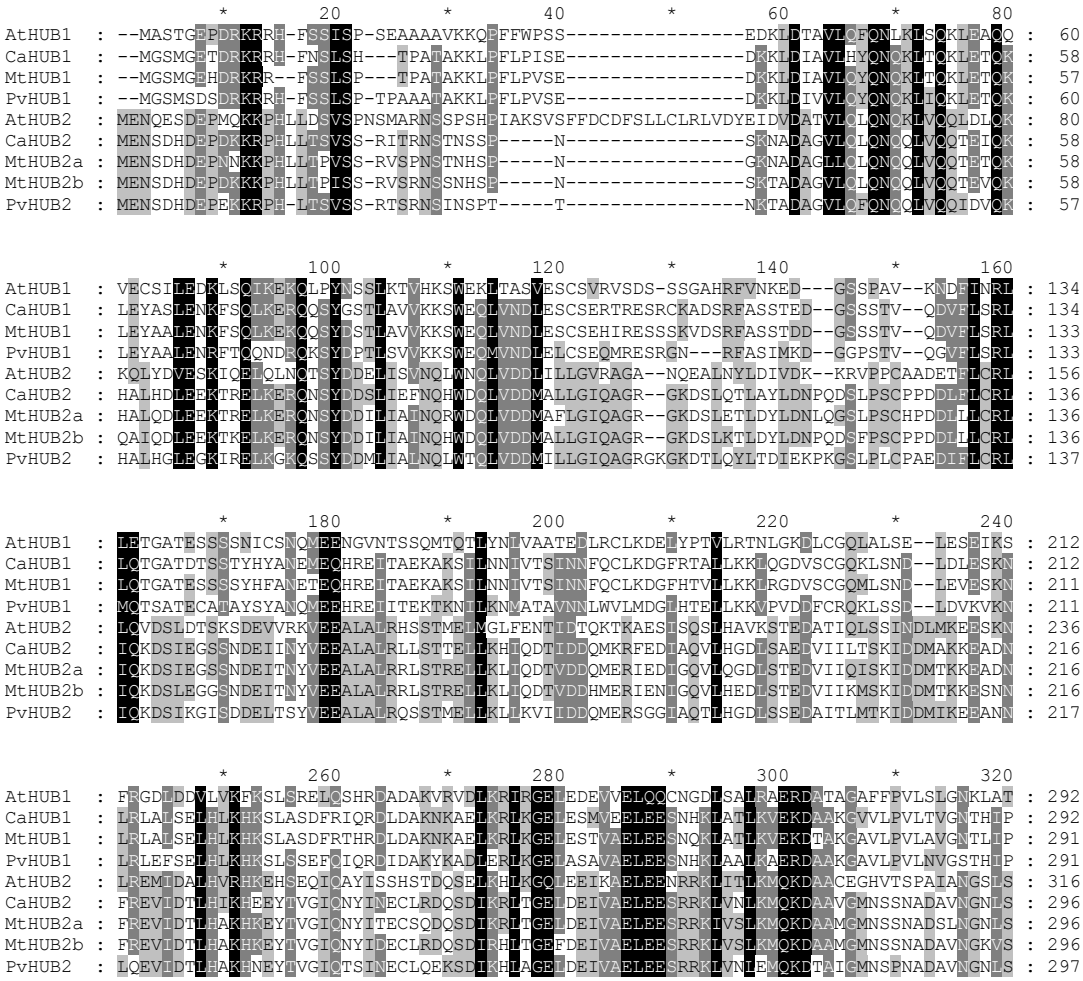
Appendix 3.37 Identity matrix (%) derived from the multiple sequence alignment of *EARLY FLOWERING 7* (ELF7) genes in the species *Arabidopsis thaliana* (At), *Phaseolus vulgaris* (Pv), *Cicer arietinum* (Ca) and *Medicago truncatula* (Mt). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

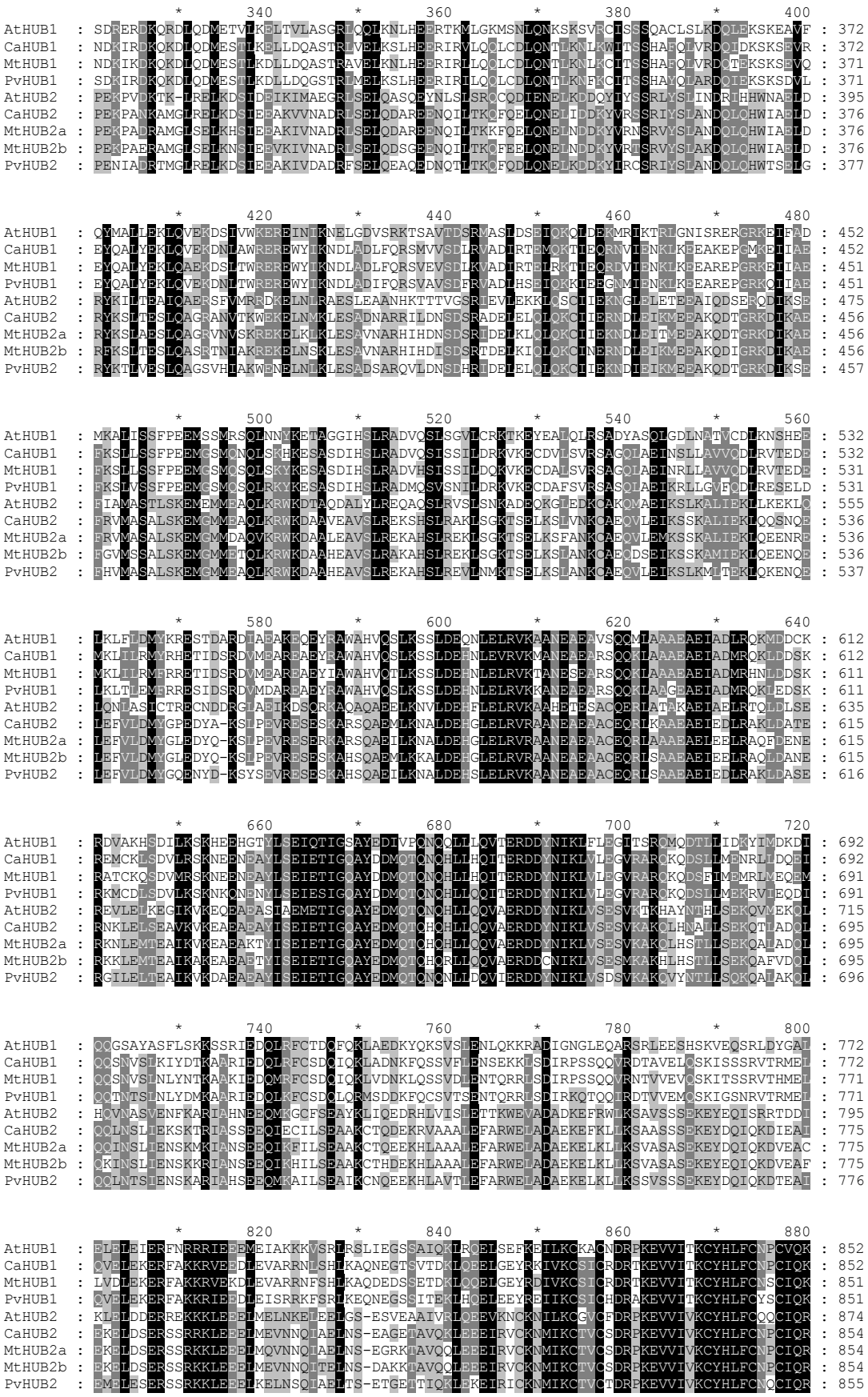
	Accession	AtELF7	PvELF7	CaELF7	MtELF7
AtELF7	AT1G79730		57.6	54.4	55.2
PvELF7	Phvul.010G145300	57.6		73.7	72.4
CaELF7	LOC101488479	54.4	73.7		81.3
MtELF7	Medtr8g014460	55.2	72.4	81.3	

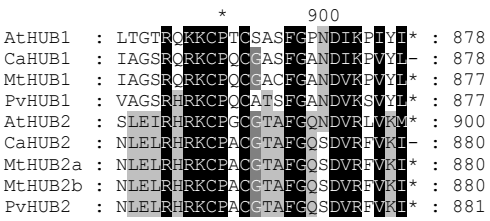


Appendix 3.38 Identity matrix (%) derived from the multiple sequence alignment of *HISTONE MONOUBIQUITINATION 1* (HUB1) and HUB2 genes in the species *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca), *Medicago truncatula* (Mt) and *Phaseolus vulgaris* (Pv). Protein sequences from accessions listed in the table were aligned using MUSCLE in Geneious 8 software.

	Acession	AtHUB1	CaHUB1	MtHUB1	PvHUB1	AtHUB2	CaHUB2	MtHUB2a	MtHUB2b	PvHUB2
AtHUB1	AT2G44950		51.1	50.5	50.2	28.6	30.7	29.4	29.4	30.2
CaHUB1	LOC101500404	51.1		87.5	75.8	30.7	34.2	32.1	31.8	33.3
MtHUB1	Medtr7g046250	50.5	87.5		74.9	30.7	33.1	31.4	31.3	33.1
PvHUB1	Phvul.008G141100	50.2	75.8	74.9		30.2	33.0	31.9	31.5	33.0
AtHUB2	AT1G55250	28.6	30.7	30.7	30.2		54.6	54.0	53.0	55.2
CaHUB2	LOC101509797	30.7	34.2	33.1	33.0	54.6		86.3	84.5	78.2
MtHUB2a	Medtr1g068830	29.4	32.1	31.4	31.9	54.0	86.3		87.9	75.5
MtHUB2b	Medtr5g085010	29.4	31.8	31.3	31.5	53.0	84.5	87.9		74.0
PvHUB2	Phvul.008G214800	30.2	33.3	33.1	33.0	55.2	78.2	75.5	74.0	




AtHUB1 : SDRERDKQRDLQDMETVLKELTVLASGRLQQIKNLHEERTKVLGKMSNLQNKSKSVECISSSQACLSLKDQLEKSKEAVF : 372
CaHUB1 : NDKIRDKQKDLQDMESILKELLDQASTRLVELKSLHEERIRVLQQLCDLQNILKNLKWITSSHAFQIVRDQIDKSKSEVR : 372
MtHUB1 : NDKIKDKQKDLQDMESILKDLLDQASTRAVELKNLHEERIRILQQLCDLQNTLKNLACITSSHAFQLVRDQTKSKSEVQ : 371
PvHUB1 : SDKIRDKQKDLQDMESILKDLLDQGSTRLMELKSLHEERIRILQQLCDLQNILKNFTCITSSHAYQLARDQIEKSKSDVL : 371
AtHUB2 : PEKPVDKTK-IRELKDSIDEIKIMAEGRLSELQASQEYNLSLSRQCQDIENELKDDQYIYSSRLYSLINDRIHHWNAELD : 395
CaHUB2 : PEKPANKAMGLRELKDSIEEAKVVNADRLSELQDAREENQILTKQFQELQNELIDDKYVRSSRIYSLANDQLQHWIAELD : 376
MtHUB2a : PEKPADRAMGLSELKHSIEEAKIVNADRLSELQDAREENQILTKKFQELQNELNDDKYVRNSRVYSLANDQLCHWIAELD : 376
MtHUB2b : PEKPAERAMGLSELKNSIEEVKIVNADRLSELQDSQREENQILTKQFEELQNELNDDKYVRLSRVYSLAKDQLQHWIAELD : 376
PvHUB2 : PENIADRTMGLRELKDSIEEAKIVDADRFSELQEAQEDNQTLTKQFQDLQNELKDDKYIRQSRIYSLANDQLQHWTSELG : 377
AtHUB1 : QVMALLEKLQVEKDSIVWKERERINIKNELGDVSRKTSAVTDSRMASLDSEIQKQLDEKMRIKTRLGNISRERGRKETEAD : 452
CaHUB1 : EVQALYEKLQVEKDNLAWRERERWYIKNDLADLFQRSMVVSDLRVADIRTEMQKTIEQRNVIENKLKEEAKEPGMKEIIAE : 452
MtHUB1 : EVQALYEKLQAEKDSLTWRERERWYIKNDLADLFQRSVEVSDLKVADIRTELRKTIEQRDVIENKLKEEAREPGRKEIIAE : 451
PvHUB1 : EVQALYEKLQVEKDNLTWRERERWYIKNDLADIFQRSVAVSLRVADIHSEIQKKIEEGNMIENKLKEEAREPGRKQIIAE : 451
AtHUB2 : RVKILTEAIQAERSFVMRRDKELNLRAESLEAANHKTTTVGSRIEVLEKKLQSCIIEKNGLPLETEEAIQDSERQDIKSE : 475
CaHUB2 : RVKSLTESLQAGRANVTKWEKELNMKLESADNARRILDNSDSRADEIELQLQKCIIERNDLEIKMEEAKQDTGRKDIKAE : 456
MtHUB2a : RVKSLAESLQAGRVNVSKREKELKLKLESAVNARHIHDNSDSRIDEIKIQLQKCIIEKNDLEITMEEAKQDTGRKDIKAE : 456
MtHUB2b : RVKSLTESLQASRTNIAKREKELNSKLESAVNARHIHDISDSRTDEIKIQLQKCINERNDLEIKMEEAKQDIGRKDIKAE : 456
PvHUB2 : RVKTLVESLQAGSVHIAKWENELNIKLESADSARQVLDNSDHRIDEIELQLQKCIIEKNDIEIKMEEAKQDTGRKDIKSE : 457
AtHUB1 : MKALISSFPEEMSSMRSQLNNVKETAGGIHSLRADVQSLSGVLCRKTKEYEALQLRSADYASQLGDLNATVCDLKNSHEE : 532
CaHUB1 : EKSLLSSFPEEMGSMQNQLSKHKESASDIHSLRADVQSISSILDRKVKECDVLSVRSAGQLAEINSLLAVVQDLRVTEDE : 532
MtHUB1 : EKSLLSSFPEEMGSMQSQLSKYKESASDIHSLRADVHSISSILDQKVKECDALSVRSAGQLAEINRLLAVVQDLRVTEDE : 531
PvHUB1 : EKSLVSSFPEEMGSMQSQLRKYKESASDIHSLRADMQSVSNILDRKVKECDAFSVRSASQLAEIKRLLGVFQDLRESELD : 531
AtHUB2 : EIAMASTLSKEMEMMEAQLKRWKDTAQDALYLREQAQSLRVSLSNKADEQKGIEDKGAKQVAEIKSLKALIEKLLKEKLQ : 555
CaHUB2 : ERVMASALSKEMGMMEAQLKRWKDAAVEAVSLREKSHSLRAKLSGKTSELKSLVNKGAEQVLEIKSSKALIEKLQQSNQE : 536
MtHUB2a : ERVMASALSKEMGMMDAQVKRWKDAALEAVSLREKAHSLREKLSGKTSELKSFANKGAEQVLEMKSSKALIEKLQEENRE : 536
MtHUB2b : EGVMSSALSKEMGMMETQLKRWKDAAHEAVSLRAKAHSLREKLSGKTSELKSLANKGAEQDSEIKSSKAMIEKLQEENQE : 536
PvHUB2 : EHVMASALSKEMGMMEAQLKRWKDAAHEAVSLREKAHSLREVLNMKTSELKSLANKGAEQVLEIKSLKMLTEKLQKENQE : 537
AtHUB1 : IKLFLDMYKRESTDARDIAEAKDQEYRAWAHVQSLKSSLDEQNLELRVKAANEAEAVSQQMIAAAEAEIADLRQKMDDCK : 612
CaHUB1 : MKLILRMYRHETIDSRDVMEAREAEYRAWAHVQSLKSSLDEFNLEVRVKMANEAEARSQQKLAAAEAEIADMRQKLDDSK : 612
MtHUB1 : MKLILRMFRRETIDSRDVMEAREAEYIAWAHVQTLKSSLDEFNLELRVKTANESEARSQQKLAAAEAEIADMRHNLDDSK : 611
PvHUB1 : IKLTLEMFRRESIDSRDVMDAREAEYRAWAHVQSLKSSLDEFNLELRVKANEAEARSQQKIAAGEAEIADMRQKLEDSK : 611
AtHUB2 : LQNLASICTRECNDDRGIAEIKDSQRKAQAQAEELKNVLDEIFLELRVKAAHETESACQERLATAKAEIAELRTQLLSE : 635
CaHUB2 : LEFVLDMYGPEDYA-KSIPEVRESESKARSQAEMLKNALDEIGLELRVRAANEAEAACEQRLKAAEAEIEDLRAKLLATE : 615
MtHUB2a : LEFVLDMYGLEDYQ-KSIPEVRESERKARSQAEILKNALDEIGLELRVRAANEAEAACEQRLAAAEAELEELRAQFEENE : 615
MtHUB2b : LEFVLDMYGLEDYQ-KSLPEVRESESKAHSQAEMLKKALDEIGLELRVRAANEAEAACEQRLSAAEAEIEELRAQLLANE : 615
PvHUB2 : LEFVLDMYGQENYD-KSYSEVRESESKAHSQAEILKNALDEISLELRVKAANEAEAACEQRLSAAEAEIEDLRAKLLASE : 616
AtHUB1 : RDVAKHSDILKSKHEEHGTYLSEIQTIGSAYEDIVPQNQQLLQVTERDDYNIKLFLEGITSRQMQDTLLIDKYIMDKDI : 692
CaHUB1 : REMCKLSDVLRSKNEENEAYLSEIETIGQAYDDMQTQNQHLLHQITERDDYNIKLVLEGVRARQKQDSLLMENRLLDQEI : 692
MtHUB1 : RATCKQSDVMRSKNEENEAYLSEIETIGQAYDDMQTQNQHLLHQITERDDYNIKLVLEGVRARQKQDSFIMEMRLVEQEM : 691
PvHUB1 : RKMCDLSDVLKSKNKQNENYLSEISIGQAYDDMQTQNQHLLQQITERDDYNIKLVLEGVRARQKQDSLLMEKRVIEQDI : 691
AtHUB2 : REVLELKEGIKVKEQEAEASIAEMETIGQAYEDMQTQNQHLLQQVAERDDYNIKLVSESVTLKHAYNTHLSEKQVVEKQL : 715
CaHUB2 : RNKLELSEAVKVKEAEAEAYISEIETIGQAYEDMQTQHQHLLQQVAERDDYNIKLVSESVKAKQLHNALLSEKQTLADQL : 695
MtHUB2a : RKNLEMTEAIKVKEAEAKTYISEIETIGQAYEDMQTQHQHLLQQVAERDDYNIKLVSESVKAKQLHSTLLSEKQALADQL : 695
MtHUB2b : RKKLEMTEAIKAKEAEAETYISEIETIGQAYEDMQTQHQRLLQQVAERDDCNIKLVSESMKAKHLHSTLLSEKQAFVDQL : 695
PvHUB2 : RGILELTEAIKVKDAEAEAYISEIETIGQAYEDMQTQNQNLLQQVIERDDYNIKLVSDSVKAKQVYNTLLSQKQALAKQL : 696
AtHUB1 : QQGSAYASFLSKKSSRIEDQLRFCTLDFQKLAEDKYQKSVSLENLQKKRADIGNGLEQARSRLEESHSKVEQSRLDYGAL : 772
CaHUB1 : QQSNVSLKIYDTKAARIEDQLRFCSDQIQKLADNKFQSSVFLENSEKKLSDIRPSSQQVRDTAVELQSKISSSRVTRMEL : 772
MtHUB1 : QQSNVSLNLYNTKAAKIEDQMRFCSDQIQKLVDNKLQSSVDLENTQRRLSDIRPSSQQVRNTVVEVQSKITSSRVTHMEL : 771
PvHUB1 : QQTNTSLNLYDMKAARIEDQLKFCSDQLQRMSDDKFQCSVTSENTQRRLSDIRKQTQQIRDTVVEMQSKIGSNRVTRMEL : 771
AtHUB2 : HQVNASVENFKARIAHNEEQMAGCFSPAYKLIQEDRHLVISLETTKWEVADAKEFRWLKSAVSSSPKEYEQISRRTDDI : 795
CaHUB2 : QQLNSLIEKSKTRIASSEEQIECILSPAAKCTQDEKRVAAALEFARWELADAEKEFKLLKSAASSSPKEYDQIQKDIEAI : 775
MtHUB2a : QQINSLIENSKMKIANSEEQIKFILSPAAKCTQEEKHLAAALEFARWELADAEKELKLLKSVASASPKEYDQIQKDVEAC : 775
MtHUB2b : QKINSLIENSKKRIANSEEQIKHILSPAAKCTHDEKHLAAALEFARWELADAEKELKLLKSVASASPKEYEQIQKDVEAF : 775
PvHUB2 : QQLNTSIENSKARIAHSEEQMKAILSPAIKCNQEEKHLAVTLEFARWELADAEKELKLLKSSVSSSPKEYDQIQKDTEAI : 776
AtHUB1 : ELELEIERFNRRRIEEEMEIAKKKVSRLRSLIEGSSAIQKLRQELSEFKEILKCKAQNDREKEVVITKCYHLFCNPCVQK : 852
CaHUB1 : QVELEKERFAKKRVEEDLEVARRNLSHLKAQNEGTSVTDKLQEELGEYRKIVKCSIGRDRTKEVVITKCYHLFCNPCIQK : 852
MtHUB1 : LVDLEKERFAKKRVEKDLEVARRNFSHLKAQDEDSSETDKLQQELGEYRDIVKCSIGRDRTKEVVITKCYHLFCNSCIQK : 851
PvHUB1 : QVELEKERFAKKRIEEDLEISRRKFSRLKEQNEGSSITEKLHQELEEYREIIKCSIGHDRAKEVVITKCYHLFCYSCIQK : 851
AtHUB2 : KLELDDERREKKKLEELMELNKELEELGS-ESVEAAIVRLQEEVKNCKNILKCGVGFDRPKEVVIKCYHLFCQQCIQR : 874
CaHUB2 : FKELDSERSSRRKLEELMEVNNQLAELNS-EAGETAVQKLEEEIRVCKNMIKCTVGSDRPKEVVIVKCYHLFCNPCIQR : 854
MtHUB2a : FKELDSERSSRKKLEELMQVNNQLAELNS-EGRKTAVQQLEEEIRVCKNMIKCTVGSDRPKEVVIVKCYHLFCNPCIQR : 854
MtHUB2b : FKELDSERSSRKKLEELMEVNNQITELNS-DAKKTAVQQLEEEIRVCKNMIKCTVGSDRPKEVVIVKCYHLFCNPCIQR : 854
PvHUB2 : EMELESERSSRKKLEELMEIKELNSQLAELTS-ETGETTIQKLEKEIRICKNMIKCTVGTDRPKEVVIVKCYHLFCNQCIQR : 855


* 900
AtHUB1 : LTGTRQKKCPTCSASFGPNDIKPIYI* : 878
CaHUB1 : IAGSRQRKCPQCGASFGANDIKPVYL- : 878
MtHUB1 : IAGSRQRKCPQCGACFGANDVKPVYL* : 877
PvHUB1 : VAGSRHRKCPQCATSFGANDVKSVYL* : 877
AtHUB2 : SLEIRHRKCPGCCTAFGQNDVRLVKM* : 900
CaHUB2 : NLELRHRKCPACCTAFGQSDVRFVKI- : 880
MtHUB2a : NLELRHRKCPACCTAFGQSDVRFVKI* : 880
MtHUB2b : NLELRHRKCPACCTAFGQSDVRFVKI* : 880
PvHUB2 : NLELRHRKCPACGTAFGQSDVRFVKI* : 881

Appendix 3.39 Maximum parsimony tree derived from the alignment of genes belonging to Polycomb Repressive Complex 2 (PRC2) in *Cicer arietinum* (Ca, in red), *Medicago truncatula* (Mt) and *Arabidopsis thaliana* (At, blue). Protein sequences from accessions listed in Table 1 were aligned using MAFFT and PAUP* software was used to build the tree. Numbers in branches represent bootstrap support from 1000 replications.

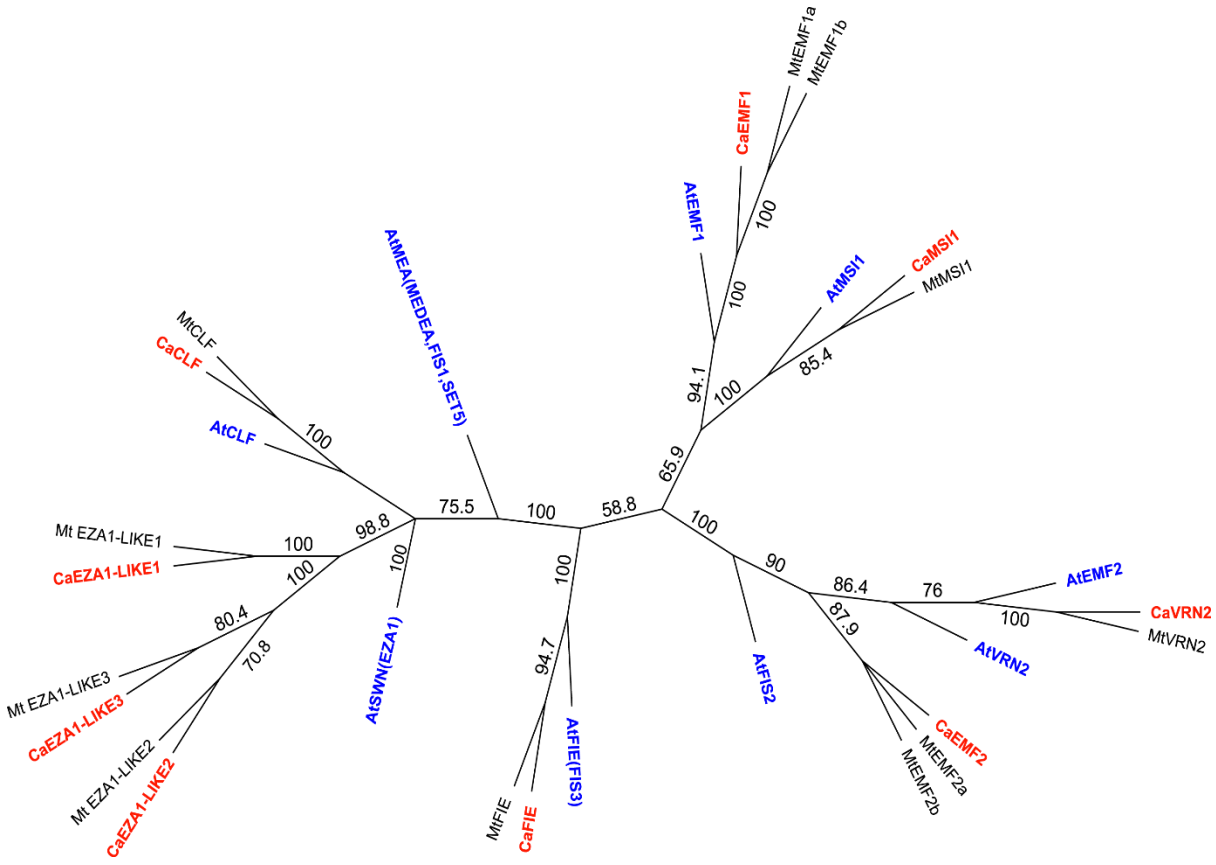


Table 1 Accession number of the sequences used in the alignment and tree construction of PRC2 genes in three plant species.

Medicago truncatula		*Arabidopsis thaliana*		*Cicer arietinum*	
CLF	Medtr5g016870	**CLF**	AT2G23380	**CLF**	LOC101514533
EMF1a	Medtr3g110082	**EMF1**	AT5G11530	**EMF1**	LOC101503745
EMF1b	Medtr3g110132	**EMF2**	AT5G51230	**EMF2**	LOC101511538
EMF2a	Medtr1g090230	**FIE(FIS3)**	AT3G20740	**FIE**	LOC101507081
EMF2b	Medtr1g090240	**FIS2**	AT2G35670	**MSI1**	LOC101515032
FIE	Medtr1g028310	**MEA (MEDEA, FIS1, SET5)**	AT1G02580	**EZA1-LIKE1**	LOC101502118
MSI1	Medtr4g096880	**MSI1**	AT5G58230	**EZA1-LIKE2**	LOC101494145
EZA1-LIKE1	Medtr1g086980	**SWN(EZA1)**	AT4G02020	**EZA1-LIKE3**	LOC101511861
EZA1-LIKE2	Medtr7g109560	**VRN2**	AT4G16845	**VRN2**	LOC101513120
EZA1-LIKE3	Medtr7g055660				
VRN2	Medtr5g013150				

Appendix 4.1 Markers used for CRIL2 genetic map, their position in the original 2956-markers map developed by the Department of Plant Pathology of University of California, Davis (Davis, USA) and the position obtained in the present study (both in cM). New markers developed for this study are highlighted in bold.

Linkage Group	Original position	Marker Name	New position
LG 1	0.000	S101p82874	83.758
	2.381	S101p642643	80.391
	4.781	S101p1155160	76.255
	6.196	S101p1468572	74.717
	9.081	S101p2152596	70.495
	10.497	S101p2533173	69.022
	13.327	S101p3863447	66.132
	16.638	S101p5217051	62.533
	21.506	S101p5614995	55.931
	26.909	S550p435913	50.101
	30.683	S550p1402959	46.264
	35.660	S1733p73076	40.812
	39.031	S129p121591	36.557
	43.607	S187p682971	32.986
	49.231	S447p573411	27.736
	53.046	S447p1498195	23.022
	56.380	S447p2884249	20.075
	60.919	S751p71909	15.027
	64.742	S751p1386566	10.491
	68.561	S751p2465070	6.069
	70.942	S751p3069740	3.664
	72.371	S751p3569871	2.080
	74.285	S751p4286627	0.000
LG 2	0.000	S521p954632	0.000
	3.805	S521p38613	3.808
	7.157	S2283p968876	7.802
	10.047	S2283p667718	10.910
	14.867	S2283p23049	16.211
	16.415	S2575p50167	18.228
	19.203	S1209p257778	20.618
	22.098	S3387p35765	23.969
	25.455	S529p23750	28.107
	28.533	S1958p111423	31.348
	30.751	S215p338782	33.529
	32.994	S270p626123	35.445
	35.966	S58p12932	38.478
	39.990	S905p486463	40.894
	43.444	S183p1679651	43.139
	45.830	S183p997009	45.919
	48.688	S183p358031	47.960
	51.541	S137p193176	52.030
	55.005	S1186p129714	56.227
	58.839	S449p289900	61.446
	62.760	S449p515924	65.677
	66.637	S449p906646	70.323
	68.996	S449p1216553	73.081

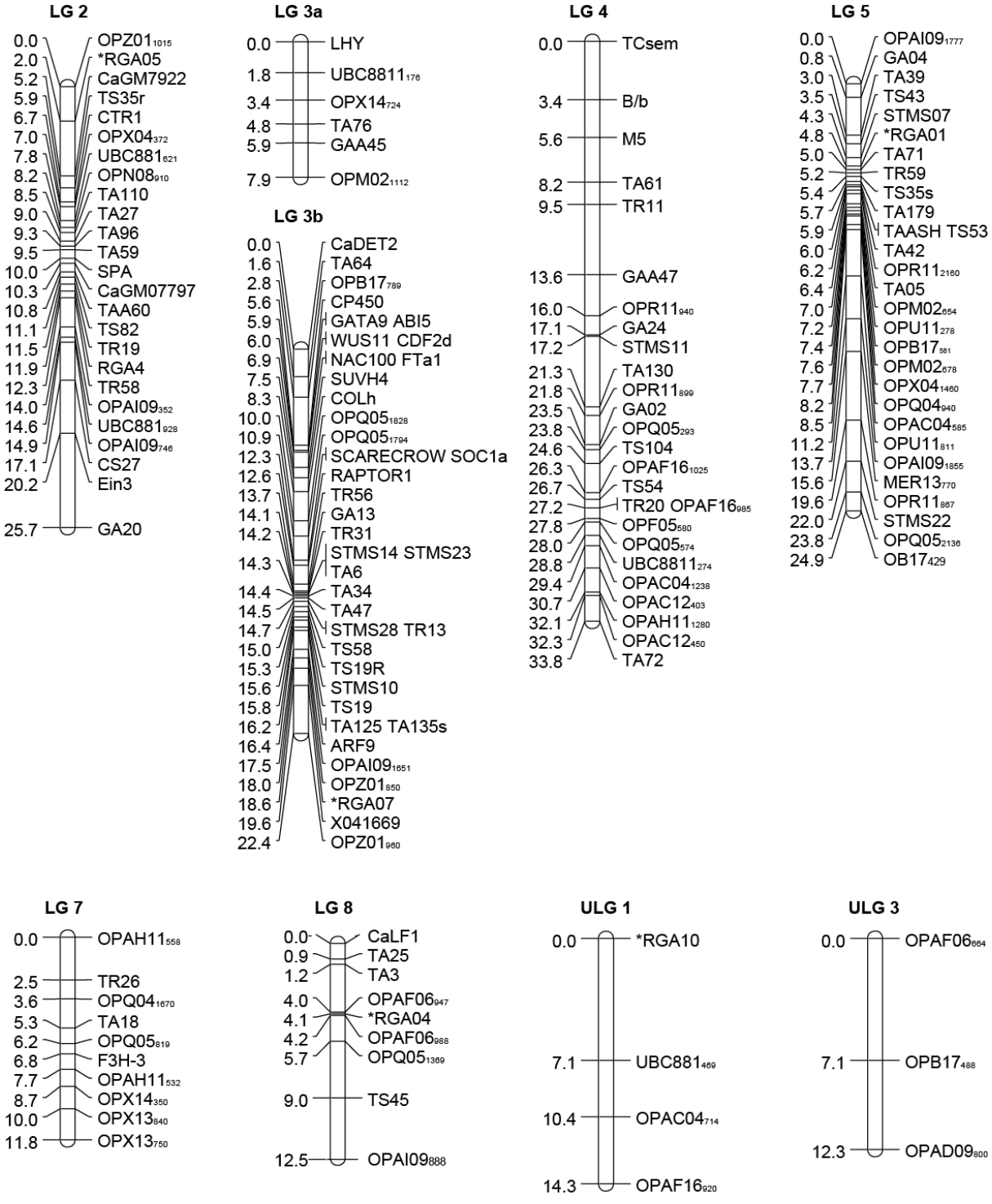
LG 3	**unknown**	***LHY***	**0.000**
	0.000	S316p5661120	0.810
	1.415	S316p5364679	1.959
	4.254	S316p4956139	4.959
	7.179	S316p4859689	8.342
	9.547	S316p3874637	10.662
	10.962	S316p3771776	12.278
	11.438	S316p3567239	12.830
	12.386	S316p3273173	14.521
	13.801	S316p2511043	16.113
	14.745	S316p2327507	16.812
	16.160	S316p1589742	18.036
	17.103	S316p1178725	19.261
	20.414	S346p384131	23.676
	23.829	S1208p282429	27.713
	24.301	S1208p166854	28.008
	27.612	S1871p1070719	32.307
	28.083	S1871p994485	32.688
	30.498	S1871p494158	35.592
	30.970	S687p462457	35.927
	31.441	S687p1009745	36.423
	31.913	S687p1454758	36.690
	32.385	S687p1508322	37.088
	32.856	S687p1864511	37.353
	unknown	**CP450**	**37.986**
	unknown	**ABI5**	**38.332**
	unknown	***CDF2d***	**38.520**
	unknown	**WUS11**	**38.520**
	unknown	**GATA9**	**39.229**
	unknown	***FTa1***	**39.462**
	unknown	**SUVH4**	**39.919**
	unknown	***COLh***	**40.003**
	34.272	S28p689334	41.029
	unknown	***SOC1a***	44.111
	unknown	**Scarecrow**	**44.654**
	37.583	S28p2932809	45.101
	unknown	**Raptor1**	**46.464**
	39.479	S714p1327693	47.357
	43.906	S1202p50545	50.305
LG 4	0.000	S34p73468	0.000
	2.418	S34p1056032	3.355
	4.323	S34p2504710	5.916
	7.177	S1978p846797	8.424
	10.479	S1758p3016775	12.433
	13.790	S1758p2397602	15.824
	18.556	S1758p608913	19.765
	22.824	S360p1277380	24.724
	26.191	S405p1034880	29.036
	29.039	S405p2263690	32.686
	32.341	S77p3281852	36.096
	35.171	S77p2647267	37.646
	37.067	S77p1658521	39.243
	41.844	S77p96279	43.922
	42.787	S2032p76148	44.818

LG 4 (continued)	45.203	S1610p104504	46.235
	48.213	S40p845447	48.764
	51.057	S431p818339	51.702
	56.369	S7355p21026	57.101
	57.785	S27p1966205	58.690
	59.200	S1534p85503	60.326
	62.048	S1534p1057661	63.816
	65.368	S1534p1913687	68.755
	67.758	S1801p378511	71.623
	70.625	S308p1196636	75.528
	72.984	S308p296417	78.295
	75.397	S308p6721	81.534
	78.320	S699p321588	84.952
LG 5	0.000	S52p738001	0.000
	3.003	S19p9163	3.286
	6.203	S188p2049560	5.147
	8.108	S188p53345	7.316
	11.450	S1301p114351	10.574
	15.975	S149p2278159	16.335
	17.862	S1596p1045628	18.456
	20.220	S660p191499	20.764
	23.087	S660p916417	23.438
	25.455	S660p2192496	25.498
	28.294	S660p3722062	29.820
	30.208	S660p3873897	31.896
	33.584	S2058p413307	35.973
	36.415	S2058p1727283	39.571
	39.245	S2058p2990262	42.719
	41.622	S2058p3537460	45.550
	43.555	S2058p3973981	46.266
	46.422	S2058p4838469	48.987
	48.826	S1211p2611133	51.773
	50.727	S1211p2033994	54.132
	54.126	S1211p1204220	57.980
	56.517	S1211p795441	60.962
	58.907	S1211p217054	63.881
LG 6	0.000	S176p117949	62.563
	2.923	S1028p430704	58.613
	4.366	S2255p617671	57.482
	7.762	S2255p362775	53.737
	11.124	S248p246648	49.862
	15.453	S248p665638	44.420
	17.811	S1280p286165	41.504
	21.645	S1030p18653	39.273
	23.060	S1030p592885	37.959
	25.423	S245p598468	36.855
	28.898	S1692p26236	34.574
	30.455	S23p627713	33.025
	30.455	S574p9146	32.829
	34.391	S976p1251259	30.157
	37.620	S2480p10422	27.455
	40.464	S1096p1969439	24.253
	43.845	S1096p715287	19.353

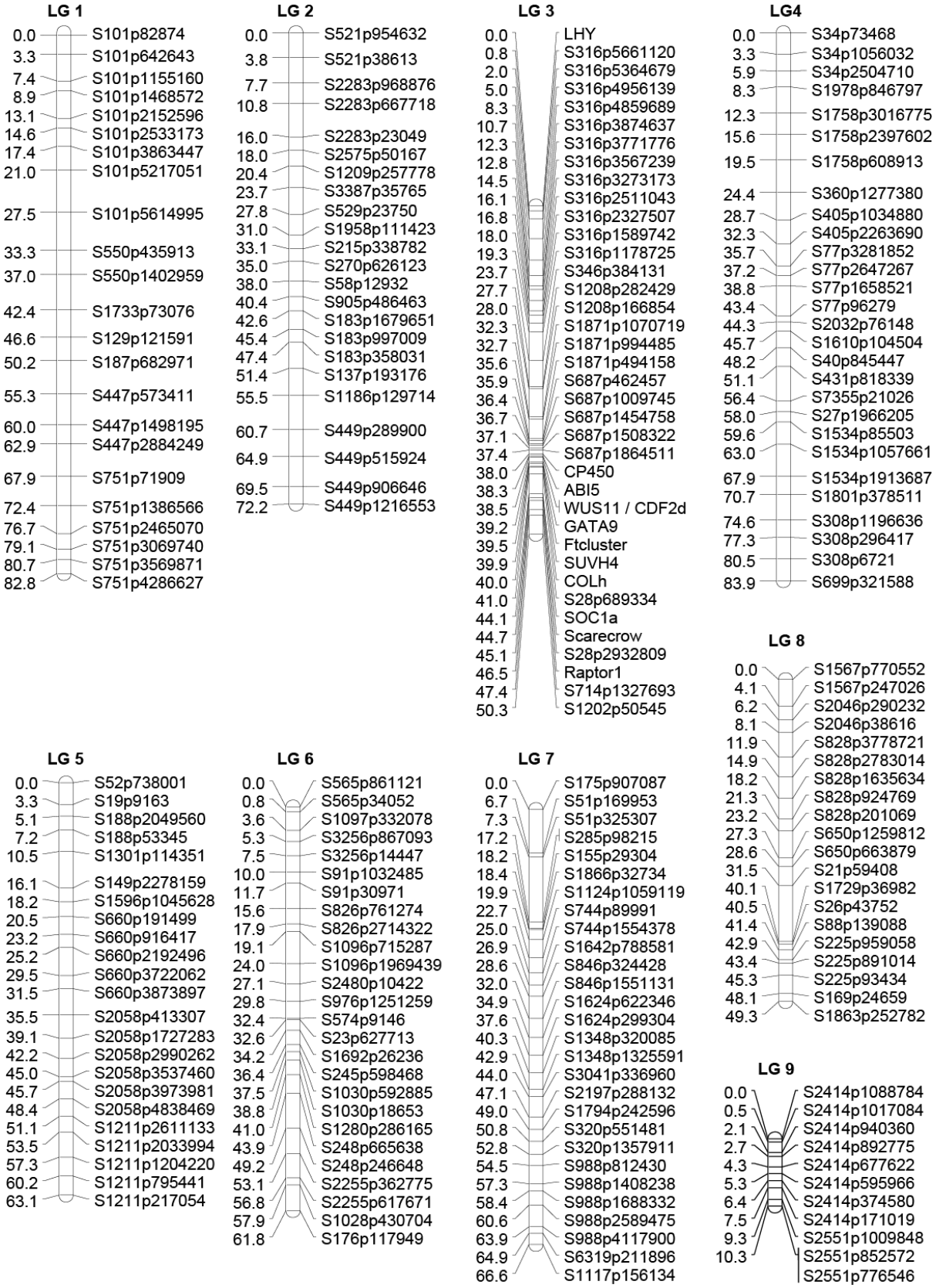
LG 6 (continued)	45.264	S826p2714322	18.094
	48.117	S826p761274	15.805
	51.932	S91p30971	11.864
	54.299	S91p1032485	10.113
	57.157	S3256p14447	7.543
	59.520	S3256p867093	5.340
	61.420	S1097p332078	3.679
	63.838	S565p34052	0.806
	64.800	S565p861121	0.000
LG 7	0.000	S175p907087	0.000
	4.549	S51p169953	6.775
	5.030	S51p325307	7.383
	13.667	S1643p95947	17.434
	13.667	S285p98215	17.430
	14.721	S155p29304	18.430
	15.206	S1866p32734	18.614
	17.121	S1124p1059119	20.170
	19.983	S744p89991	22.962
	21.870	S744p1554378	25.301
	23.765	S1642p788581	27.270
	25.661	S846p324428	28.950
	28.514	S846p1551131	32.437
	31.354	S1624p622346	35.282
	33.712	S1624p299304	38.089
	36.094	S1348p320085	40.847
	38.461	S1348p1325591	43.432
	39.418	S3041p336960	44.580
	42.452	S2197p288132	47.661
	44.344	S1794p242596	49.627
	46.258	S320p551481	51.448
	48.162	S320p1357911	53.452
	50.557	S988p812430	55.194
	52.732	S988p1408238	57.962
	54.170	S988p1688332	59.167
	56.088	S988p2589475	61.384
	58.531	S988p4117900	64.708
	59.750	S6319p211896	65.727
	61.193	S1117p156134	67.372
LG 8	0.000	S1567p770552	49.873
	3.362	S1567p247026	45.775
	5.729	S2046p290232	43.620
	7.662	S2046p38616	41.676
	10.964	S828p3778721	37.825
	13.323	S828p2783014	34.850
	16.625	S828p1635634	31.474
	19.455	S828p924769	28.323
	21.369	S828p201069	26.427
	24.721	S650p1259812	22.259
	26.150	S650p663879	20.939
	28.540	S21p59408	17.994
	36.858	S1729p36982	9.234
	37.426	S26p43752	8.902
	39.781	S88p139088	7.908

LG 8 (continued)	41.667	S225p959058	6.448
	42.139	S225p891014	5.915
	44.044	S225p93434	4.015
	46.412	S169p24659	1.162
	47.364	S1863p252782	0.000
LG 9	0.000	S2551p852572	0.000
	0.000	S2551p776546	0.000
	1.443	S2551p1009848	0.998
	3.385	S2414p171019	2.867
	4.338	S2414p374580	3.942
	5.281	S2414p595966	5.025
	6.234	S2414p677622	6.106
	7.182	S2414p892775	7.726
	7.658	S2414p940360	8.261
	8.620	S2414p1017084	9.865
	9.096	S2414p1088784	10.402

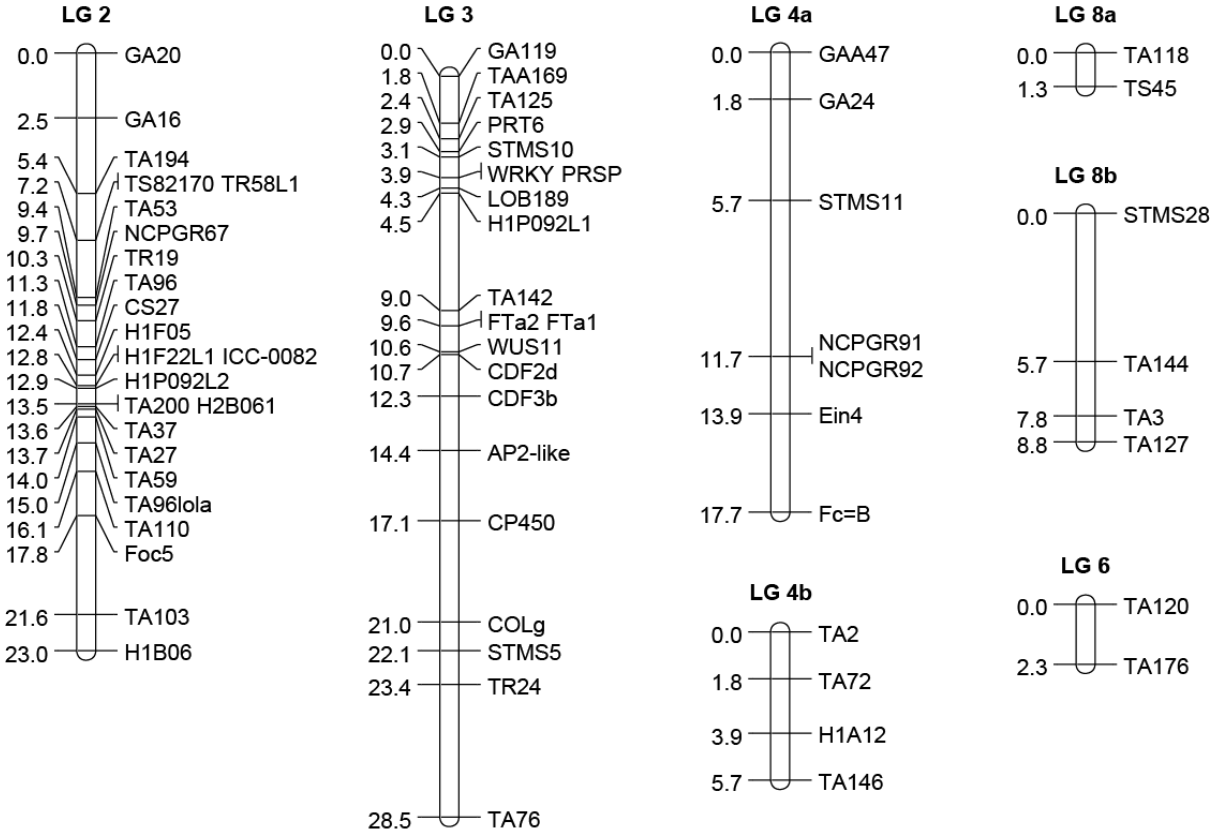
Appendix 4.2 Genetic map constructed from RIP12 (ICCL81001 x Cr5-9). Numbers in the left represent distances in cM.



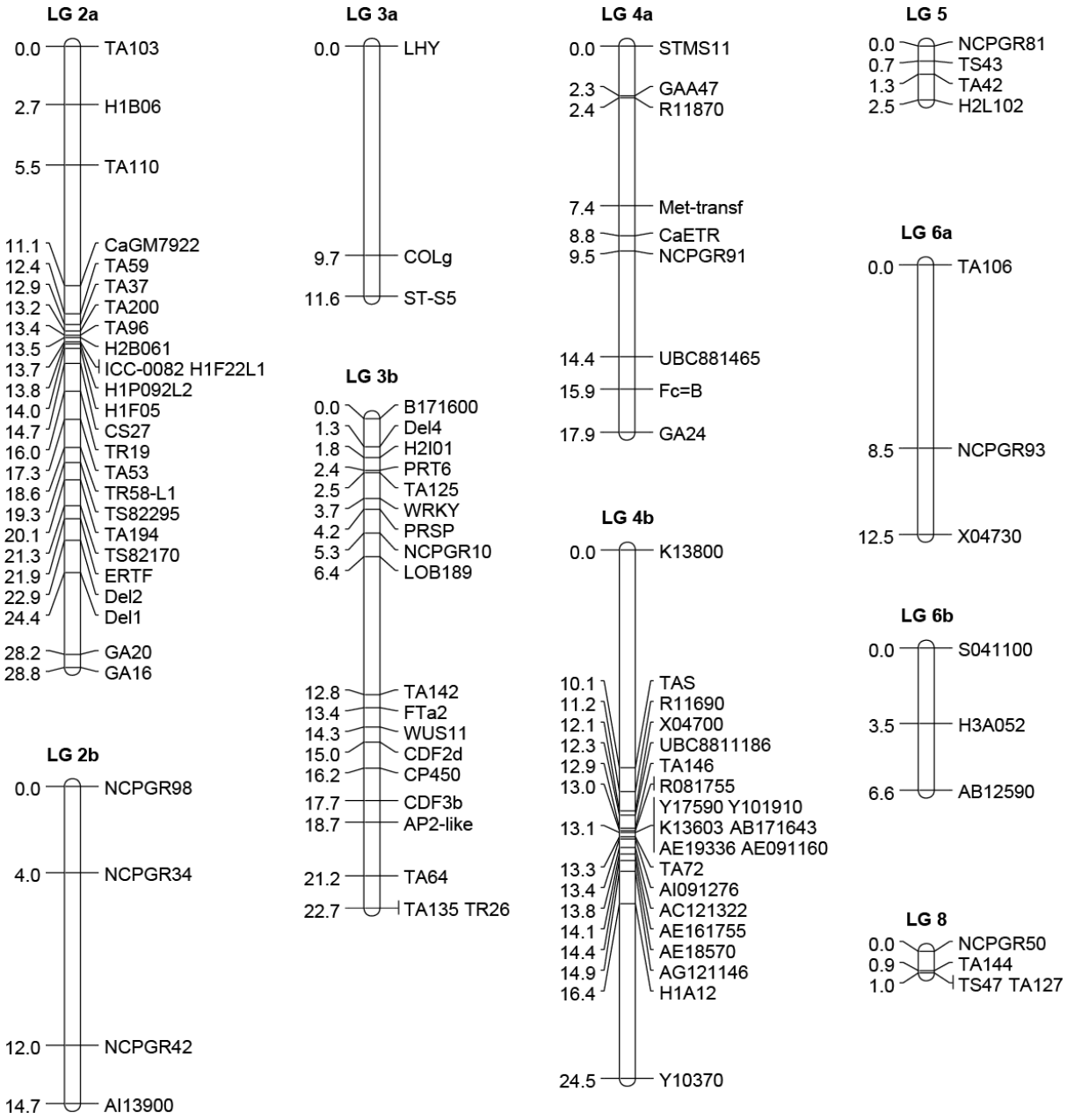
Appendix 4.3 Genetic map constructed from CRIL2 (ICC4958 x PI489777). Numbers in the left represent distances in cM.



Appendix 4.4 Genetic map constructed from RIP5 (WR315 x ILC3279). Numbers in the left represent distances in cM.



Appendix 4.5 Genetic map constructed from RIP8 (ILC3279 x WR315). Numbers in the left represent distances in cM.



Appendix 4.6 Gene ID, start position of the gene in chickpea chromosome 3 and description of the 122 genes existing between markers SUVH4 and CDF2d according to NCBI.

Marker	Position	GeneID	Description
SUVH4	25748486	101508428	histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH4
	25765688	101509380	CLAVATA3/ESR (CLE)-related protein 4A-2-like
	25787922	101509702	uncharacterized protein LOC101509702
	25849917	101510346	protein Mpv17-like, partial
	25899924	101510883	N-alpha-acetyltransferase MAK3
	25908342	101511521	cytochrome b-c1 complex subunit Rieske-4, mitochondrial-like
	25911181	101511199	uncharacterized protein LOC101511199
	25920520	101511847	uncharacterized protein LOC101511847
	25929005	101512385	uncharacterized protein LOC101512385
	25945170	101512699	LRR receptor-like serine/threonine-protein kinase RPK2
	25987693	101513246	UPF0496 protein At3g19330-like
	26006715	101513575	putative serine/threonine-protein kinase isoform X2
	26029555	101514543	chalcone synthase
	26038727	101505971	uncharacterized protein LOC101505971
	26049105	101514869	jasmonate O-methyltransferase-like isoform X2
	26056836	101506278	jasmonate O-methyltransferase-like
	26073603	101506600	3,7-dimethylxanthine N-methyltransferase-like
	26105741	101515408	7-methylxanthosine synthase 1-like
	26119620	101506937	ethylene-responsive transcription factor LEP
	26135351	101515744	nudix hydrolase 12, mitochondrial
	26162323	101488713	calcium-binding mitochondrial carrier protein SCaMC-1-like
	26169052	101489051	uncharacterized protein LOC101489051
	26187333	101489383	histidine biosynthesis bifunctional protein hisIE, chloroplastic
	26191225	101489915	uncharacterized protein LOC101489915
	26198363	105851804	glycine-rich cell wall structural protein 1-like
	26211557	101490249	syntaxin-22-like
	26221867	101490568	syntaxin-22-like
	26223872	101490780	uncharacterized protein LOC101490780
	26228217	101491110	DNA cross-link repair protein SNM1 isoform X1
	26233285	101491643	uncharacterized protein LOC101491643
	26237155	101491949	uncharacterized protein LOC101491949
	26259086	101492622	uncharacterized protein LOC101492622

Appendix 4.6 Continued

	26268535	101492953	probable xyloglucan endotransglucosylase/hydrolase protein B
	26275124	101493288	probable xyloglucan endotransglucosylase/hydrolase-like precursor
	26293522	101493610	glyceraldehyde-3-phosphate dehydrogenase A, chloroplastic
	26300947	101493929	pentatricopeptide repeat-containing protein At3g26630, chloroplastic
	26304938	101494250	WD repeat-containing protein LWD1
	26309206	105851806	pentatricopeptide repeat-containing protein At3g26630, chloroplastic-like
	26315891	101494555	acetyl-coenzyme A synthetase, chloroplastic/glyoxysomal
	26325185	101494873	eukaryotic peptide chain release factor subunit 1-3
	26331419	101495197	importin-13 isoform X2
	26353855	101495954	INO80 complex subunit D-like
	26365580	101496291	omega-hydroxypalmitate O-feruloyl transferase
FTa1	26393854	101497376	protein FLOWERING LOCUS T-like
	26409508	101496618	protein HEADING DATE 3A-like isoform X1
	26437711	101508200	protein FLOWERING LOCUS T-like
	26446273	101497706	transmembrane and coiled-coil domain-containing protein 1-like
	26454060	101498244	aquaporin SIP1-2-like
	26459034	101498578	apyrase 2-like
	26465157	101499141	pentatricopeptide repeat-containing protein At5g15010, mitochondrial
	26476855	101499461	nucleoside-triphosphatase-like
	26505866	101499961	NAC transcription factor 29
	26513981	101500308	uncharacterized protein LOC101500308
NAC100	26542626	101500623	NAC domain-containing protein 100-like
	26551883	101500931	U11/U12 small nuclear ribonucleoprotein 65 kDa protein
	26560819	101501672	thioredoxin-like 1-1, chloroplastic
	26568175	101501985	CCR4-NOT transcription complex subunit 3 isoform X1
	26586013	101503370	glucan endo-1,3-beta-glucosidase 8-like
GATA9	26593826	101503040	GATA transcription factor 9-like
	26551883	101500931	U11/U12 small nuclear ribonucleoprotein 65 kDa protein
	26560819	101501672	thioredoxin-like 1-1, chloroplastic
	26568175	101501985	CCR4-NOT transcription complex subunit 3 isoform X1
	26586013	101503370	glucan endo-1,3-beta-glucosidase 8-like
	26593826	101503040	GATA transcription factor 9-like
	26602473	101502729	uncharacterized protein LOC101502729
	26609037	101503708	PHD finger protein ING2
	26615005	101504033	photosynthetic NDH subunit of subcomplex B 2, chloroplastic

Appendix 4.6 Continued

	26622693	101504344	uncharacterized protein LOC101504344 isoform X1
	26633172	101504885	uncharacterized protein LOC101504885
	26645639	101505413	inosine-5\'-monophosphate dehydrogenase-like
	26653025	101508837	E3 ubiquitin-protein ligase SINA-like 4
	26659004	101505747	dnaJ homolog subfamily B member 7
	26662326	101506066	uncharacterized protein LOC101506066
	26676554	101506382	glutamate dehydrogenase 1
	26680912	101507162	putative H/ACA ribonucleoprotein complex subunit 1-like protein 1
	26688782	101507678	adenylyl-sulfate kinase 3
	26695494	101508429	uncharacterized protein LOC101508429
	26702495	101508950	LOW QUALITY PROTEIN: uncharacterized protein LOC101508950
	26706078	101509159	serine/threonine-protein phosphatase 7 long form homolog
	26720873	101509263	arabinogalactan peptide 14-like
	26725442	101509579	40S ribosomal protein S20-2-like
	26732400	101509906	uncharacterized protein LOC101509906
	26737749	101510225	tubulin alpha-3 chain-like
	26742766	101510560	ethylene-responsive transcription factor RAP2-11-like
	26748685	101510884	importin-5-like
	26768276	101511409	probable glycosyltransferase At5g03795
	26773933	101511741	DNA ligase 1
	26778850	101512057	cation/H(+) antiporter 15-like
	26823821	101513347	spermidine hydroxycinnamoyl transferase-like
	26827349	101513669	spermidine hydroxycinnamoyl transferase-like
	26830225	101513999	lysosomal beta glucosidase-like isoform X2
	26844188	101515516	uncharacterized protein LOC101515516
	26868801	101488266	binding partner of ACD11 1
	26877749	101488844	mavicyanin-like
	26879612	101489166	ALA-interacting subunit 3-like
	26886956	101489498	inactive protein RESTRICTED TEV MOVEMENT 2-like
	26888197	101489815	mavicyanin-like
	26891715	101490138	mavicyanin-like
	26895715	101510448	mavicyanin-like
	26899605	101490461	nucleolar MIF4G domain-containing protein 1
	26908441	101490782	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase-like
	26932779	101491644	keratin, type I cytoskeletal 9-like

Appendix 4.6 Continued

	26942082	101511105	uncharacterized protein LOC101511105
	26948508	101491950	protein LHCP TRANSLOCATION DEFECT
	26952798	101511410	uncharacterized protein LOC101511410
	26955920	101492284	pentatricopeptide repeat-containing protein At4g39530
	26965238	101492623	U5 small nuclear ribonucleoprotein 40 kDa protein
	26968496	101492954	exosome component 10-like isoform X2
	26980897	101493492	plant cysteine oxidase 2-like
	26986069	101493811	diphthine--ammonia ligase
	27002842	101494557	probable inactive leucine-rich repeat receptor-like protein kinase At3g03770
	27009392	101495088	probable zinc metalloprotease EGY1, chloroplastic
	27019556	101495418	uncharacterized protein LOC101495418 isoform X1
	27028135	101495749	uncharacterized protein LOC101495749
	27041619	101496072	BTB/POZ and MATH domain-containing protein 4
	27056650	101496400	cationic amino acid transporter 4, vacuolar
	27066225	101496734	cationic amino acid transporter 2, vacuolar-like
	27073720	101497042	putative disease resistance RPP13-like protein 1
	27085215	101497377	LOW QUALITY PROTEIN: probable polyribonucleotide nucleotidyltransferase 1, chloroplastic
	27104961	101497708	uncharacterized protein LOC101497708
	27108659	101498031	cation/calcium exchanger 1
	27122380	101498362	cation/calcium exchanger 2-like
	27132304	101498703	uncharacterized protein LOC101498703
	27135177	101499027	protein PNS1-like
	27143886	101499344	xylosyltransferase 2-like isoform X1
	27150286	101511742	transcription factor IIIB 90 kDa subunit-like
	27157582	101500084	transcription factor IIIB 90 kDa subunit-like
	27177500	101500413	uncharacterized protein LOC101500413
CDF2d	27192177	101500722	cyclic dof factor 3-like

Appendix 4.7 Gene ID, start position of the gene in chickpea chromosome 3 and description of the 244 genes existing between markers PRT6 and LOB189, according to NCBI. Interesting genes potentially involved in flowering control are highlighted in red.

Position	GeneID	Description
17915563	101506928	E3 ubiquitin-protein ligase PRT6
17936072	101507474	transcription factor AS1
17984014	101508197	probable serine/threonine-protein kinase NAK
18005002	101508735	expansin-A4-like
18034418	101509053	inactive rhomboid protein 1-like
18048686	101509376	oxysterol-binding protein-related protein 3A-like
18078534	101509698	telomere repeat-binding protein 6-like
18109283	101510018	scarecrow-like protein 9
18114370	101510337	uncharacterized protein At5g02240-like
18119889	101511943	protein FAR1-RELATED SEQUENCE 6-like
18174387	101512605	uncharacterized protein LOC101512605
18176880	101512912	uncharacterized protein LOC101512912
18177498	105851775	uncharacterized protein LOC105851775
18178070	101513238	serine/threonine-protein phosphatase 7 long form homolog
18214027	101513566	protein NRT1/ PTR FAMILY 5.6-like
18262310	101513876	probable proteasome inhibitor
18286999	101514220	uncharacterized protein LOC101514220
18303320	101495304	uncharacterized protein LOC101495304
18311955	101514538	uncharacterized protein LOC101514538
18330255	101514862	putative lipid-transfer protein DIR1
18336160	101515189	rop guanine nucleotide exchange factor 7-like
18355079	101488595	uncharacterized protein LOC101488595
18371441	101488259	probable glutathione S-transferase
18374536	101489906	probable glutathione S-transferase
18376294	101489588	uncharacterized protein LOC101489588
18381223	105851783	probable glutathione S-transferase
18385922	101488943	probable glutathione S-transferase
18388998	101490242	uncharacterized protein C167.05-like
18497412	101491206	copper transporter 6-like
18511934	101491519	copper transporter 6-like
18515549	101491829	18.1 kDa class I heat shock protein-like
18538849	101492161	uncharacterized protein LOC101492161
18555801	101492499	phosphatidylinositol:ceramide inositolphosphotransferase 2 isoform X2
18572400	101493065	uncharacterized protein LOC101493065
18616599	105851779	uncharacterized protein LOC105851779
18642474	101493388	laccase-7-like
18653887	101493711	F-box/kelch-repeat protein At3g23880-like

Appendix 4.7 Continued

18686452	101494346	fructose-1,6-bisphosphatase, chloroplastic
18730118	101494658	L-type lectin-domain containing receptor kinase VIII.2
18734377	101494979	DNA-directed RNA polymerases II, IV and V subunit 12
18739322	105851780	uncharacterized protein LOC105851780
18795544	105851781	uncharacterized protein LOC105851781
18822195	101498574	uncharacterized protein LOC101498574 isoform X1
18828620	101500186	uncharacterized protein LOC101500186
18832136	101500501	putative zinc transporter At3g08650
18853980	101500823	putative UPF0481 protein At3g02645
18888263	101488596	uncharacterized protein LOC101488596
18889420	101488945	serine/threonine-protein phosphatase 7 long form homolog
18919746	101501146	uncharacterized protein LOC101501146 isoform X4
18957227	101501452	uncharacterized protein LOC101501452 isoform X1
19015046	101502942	uncharacterized protein LOC101502942
19044301	101490244	pentatricopeptide repeat-containing protein At3g09060
19046983	101503271	uncharacterized protein LOC101503271
19048813	101503592	protein NRT1/ PTR FAMILY 8.1-like
19075719	101504133	transcription factor DIVARICATA-like
19086800	101490564	uncharacterized protein LOC101490564
19102196	101504449	sulfoquinovosyl transferase SQD2-like
19119014	101504766	protein SUPPRESSOR OF npr1-1, CONSTITUTIVE 1-like
19126695	101505087	uncharacterized protein LOC101505087
19151800	101490887	ABC transporter A family member 7-like
19174266	101505408	glycerol-3-phosphate 2-O-acyltransferase 6-like
19271074	101505743	auxin transporter-like protein 2
19283317	101506710	T-complex protein 1 subunit beta isoform X2
19283457	101506063	T-complex protein 1 subunit beta-like isoform X2
19300882	101507256	uncharacterized protein At4g19900-like
19336671	101507570	RNA polymerase II C-terminal domain phosphatase-like 2
19350387	101507884	putative serine/threonine-protein kinase isoform X2
19355856	101508421	uncharacterized protein LOC101508421
19389652	101508942	mucin-17-like
19409301	101509258	protein yippee-like At3g08990 isoform X1
19421482	101509802	high mobility group B protein 1
19426146	101510339	uncharacterized protein LOC101510339
19435745	101510878	protein root UVB sensitive 5 isoform X1
19465698	101511406	alcohol dehydrogenase-like 4
19474728	101511944	probable envelope ADP, ATP carrier protein, chloroplastic
19488536	101512695	vacuolar cation/proton exchanger 3
19522700	101513239	LIMR family protein At5g01460

Appendix 4.7 Continued

19535435	101513567	GDSL esterase/lipase CPRD49-like
19564152	101513877	GDSL esterase/lipase CPRD49-like
19573363	101514221	GDSL esterase/lipase CPRD49-like
19582955	101514539	GDSL esterase/lipase CPRD49-like
19599890	101514863	GDSL esterase/lipase CPRD49-like isoform X1
19653247	101515190	GDSL esterase/lipase CPRD49-like
19712301	101515511	GDSL esterase/lipase CPRD49-like
19724518	101488261	GDSL esterase/lipase CPRD49-like
19780841	101491520	protein trichome birefringence-like 3 isoform X1
19785923	101492059	uncharacterized protein LOC101492059
19791553	101492400	protein SCO1 homolog 1, mitochondrial
19797060	101492733	AMP deaminase-like
19818020	101493284	TMV resistance protein N-like isoform X2
19833919	105851785	uncharacterized protein LOC105851785
19845351	101495080	chlorophyll synthase, chloroplastic
19850475	101495630	protein SLE1
19854286	101495952	ERBB-3 BINDING PROTEIN 1
19863805	101496288	importin-11
19903945	101496837	dnaJ protein ERDJ3A
19922401	101497155	pyruvate decarboxylase 2
19950586	101497699	BAG family molecular chaperone regulator 4 isoform X2
19958640	101488385	uncharacterized protein LOC101488385
19975339	101498240	ammonium transporter 2
19985012	101498576	putative Myb family transcription factor At1g14600 isoform X2
19997213	101488710	serine/threonine-protein phosphatase 7 long form homolog
20000189	101499134	uncharacterized protein LOC101499134
20000854	101489046	uncharacterized protein LOC101489046
20006901	101499454	uncharacterized protein LOC101499454
20010350	101499759	uncharacterized protein LOC101499759 isoform X1
20020793	101500077	ethylene-overproduction protein 1
20029921	101500408	uncharacterized protein LOC101500408
20036107	101500718	nodulation-signaling pathway 2 protein-like
20059068	101501030	uncharacterized protein LOC101501030
20063603	101489378	uncharacterized protein LOC101489378
20069127	101501348	uncharacterized protein LOC101501348
20072333	101501670	uncharacterized protein LOC101501670
20080159	101501983	uncharacterized protein LOC101501983
20089388	101502524	mitotic-spindle organizing protein 1B-like
20092481	101502831	40S ribosomal protein S4-2-like
20098769	101503152	ferritin-1, chloroplastic-like isoform X2

Appendix 4.7 Continued

20102451	101503473	uncharacterized protein At5g01610-like, partial
20132214	101503815	abscisic acid receptor PYL4
20151167	101504134	protein trichome birefringence-like 34
20155446	101504673	proteasome subunit alpha type-4-like
20170857	101504994	probable inactive leucine-rich repeat receptor-like protein kinase At1g66830
20176048	101505307	MATE efflux family protein 2, chloroplastic
20186228	101505842	uncharacterized protein LOC101505842
20203278	105851792	uncharacterized protein LOC105851792
20210352	101506164	40S ribosomal protein S11
20239073	101506494	nucleolar complex protein 3 homolog
20275064	101506823	macrophage migration inhibitory factor homolog
20282695	101490361	protein SIEL
20292515	101507160	uncharacterized protein LOC101507160 isoform X3
20328505	101507675	dirigent protein 19-like
20330838	101508000	pentatricopeptide repeat-containing protein At3g18020
20362350	101508309	patellin-6-like
20383954	101508626	aldose reductase
20387144	101508944	probable protein phosphatase 2C 65
20418594	101509260	uncharacterized protein LOC101509260
20462194	101509578	soyasapogenol B glucuronide galactosyltransferase
20471104	101509904	soyasapogenol B glucuronide galactosyltransferase
20476492	101510223	soyasapogenol B glucuronide galactosyltransferase
20487276	105851787	uncharacterized protein LOC105851787
20490048	101510559	soyasapogenol B glucuronide galactosyltransferase-like
20495289	101510880	soyasapogenol B glucuronide galactosyltransferase
20503215	101511198	CASP-like protein 4B1
20506092	101511519	probable WRKY transcription factor 33
20527043	101511844	putative GDP-L-fucose synthase 2
20533781	101512157	uncharacterized protein LOC101512157 isoform X1
20548140	101512606	TMV resistance protein N
20552049	101512913	protein SCAR3-like
20607324	101513240	short-chain dehydrogenase reductase 2a-like
20668228	101513568	wound-induced protein 1-like
20671728	101513878	uncharacterized protein LOC101513878
20686896	101514222	alpha-soluble NSF attachment protein 2-like isoform X1
20703116	101492280	uncharacterized protein LOC101492280
20773997	105851078	protein LURP-one-related 10-like
20778319	101492951	protein LURP-one-related 15-like
20803945	101493606	pentatricopeptide repeat-containing protein At2g38420, mitochondrial
20818473	101514540	class E vacuolar protein-sorting machinery protein hse1-like

Appendix 4.7 Continued

20828026	101514864	regulatory-associated protein of TOR 1-like isoform X2
20860993	101515404	diaminopimelate epimerase, chloroplastic-like
20866547	101515739	E3 ubiquitin-protein ligase HERC2
20879217	101488709	alanine--glyoxylate aminotransferase 2 homolog 3, mitochondrial-like
20912504	101494870	uncharacterized protein LOC101494870
20930665	101495195	uncharacterized protein LOC101495195
20941722	101489270	G-type lectin S-receptor-like serine/threonine-protein kinase SD2-5 isoform X1
20950435	101489590	transcription factor SCREAM2-like
20964550	101489909	cationic peroxidase 1-like
20969188	101490245	cationic peroxidase 1-like
20981256	101490566	stress response protein NST1-like
21005024	101490888	putative uncharacterized protein DDB_G0277255
21019172	101491209	WEB family protein At2g38370-like isoform X1
21029971	101491522	prosaposin
21045152	101492060	CBL-interacting protein kinase 18-like
21049508	101492401	uncharacterized protein LOC101492401
21068003	101492734	lysine--tRNA ligase-like
21100017	101493164	CBL-interacting serine/threonine-protein kinase 14-like
21133271	101493488	transcription repressor OFP2-like
21155928	101493807	uncharacterized protein LOC101493807
21162372	101494133	probable serine/threonine-protein kinase WNK5
21173671	101494440	uncharacterized protein LOC101494440
21184453	101495519	serine/threonine-protein kinase HT1 isoform X1
21188617	101494754	non-specific lipid-transfer protein 1-like
21198377	101495082	non-specific lipid-transfer protein 1-like
21209814	101495412	non-specific lipid-transfer protein 1-like
21218454	101495742	RING-H2 finger protein ATL74-like
21230819	101496065	probable receptor-like protein kinase At3g55450
21234105	101496174	B3 domain-containing transcription repressor VAL2-like
21244578	101496395	uncharacterized protein LOC101496395
21251583	101496727	pentatricopeptide repeat-containing protein At5g41170, mitochondrial-like
21264430	101497036	uncharacterized protein LOC101497036 isoform X1
21277660	101497589	leucine-rich repeat receptor-like protein kinase PXC2
21292196	101497923	protein FATTY ACID EXPORT 3, chloroplastic isoform X1
21301452	101496504	uncharacterized protein LOC101496504
21306189	101498912	vacuolar-sorting receptor 1-like
21334539	101499956	uncharacterized protein LOC101499956
21365982	101500302	probable WRKY transcription factor 70
21379075	101500614	phospholipid-transporting ATPase 1-like
21432419	101500924	high mobility group B protein 6

Appendix 4.7 Continued

21434993	101501255	uncharacterized protein LOC101501255
21443513	101501564	serine/threonine-protein kinase AtPK2/AtPK19-like
21447428	105851788	serine/threonine-protein kinase STN8, chloroplastic-like
21448878	101502423	serine/threonine-protein kinase STN8, chloroplastic-like
21453843	101502100	mannan endo-1,4-beta-mannosidase 6
21458983	101503593	alpha-galactosidase 1
21468412	101503915	receptor-like protein kinase FERONIA
21474662	101504235	thioredoxin H-type-like
21486062	101504767	ubiquitin-conjugating enzyme E2 28
21492553	101505088	coiled-coil domain-containing protein 39 isoform X1
21499739	101505409	probable inactive receptor kinase At5g58300
21506550	101505745	DNA-directed RNA polymerase II subunit 1
21521517	101506064	4-hydroxybenzoate polyprenyltransferase, mitochondrial
21530601	101507044	BTB/POZ domain-containing protein At3g08570-like
21537552	101507363	probable acetyltransferase NSI
21547510	101507676	probable LRR receptor-like serine/threonine-protein kinase At1g06840
21558269	101508001	pentatricopeptide repeat-containing protein At1g56690, mitochondrial-like
21594431	101508311	KH domain-containing protein At2g38610-like
21608004	101508628	uncharacterized protein LOC101508628
21627330	101508946	RING finger and transmembrane domain-containing protein 2
21633471	101509484	uncharacterized protein LOC101509484
21640963	101510020	uncharacterized protein LOC101510020
21646444	101510340	pre-mRNA-splicing factor ISY1 homolog
21654218	101510671	uncharacterized protein LOC101510671
21661625	105851789	uncharacterized protein LOC105851789
21665790	101510992	calcium-binding allergen Bet v 3-like
21670690	101511300	uncharacterized protein LOC101511300
21673438	101511623	uncharacterized protein LOC101511623
21677367	101511945	mediator of RNA polymerase II transcription subunit 33A-like
21690174	105851790	uncharacterized protein LOC105851790
21696832	101512486	uncharacterized protein LOC101512486
21715046	101512802	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
21732536	101513130	serine/threonine-protein kinase At5g01020
21736501	101513456	ethanolamine-phosphate cytidylyltransferase
21773761	101513767	protein SCARECROW
21789334	101497487	probable galacturonosyltransferase 7
21802507	101514108	probable protein S-acyltransferase 14
21816861	101514431	vestitone reductase
21821573	101514980	vestitone reductase-like
21828212	101515305	vestitone reductase-like

Appendix 4.7 Continued

21830345	101497823	vestitone reductase-like
21858278	101498818	transducin beta-like protein 2
21866535	101499136	methylsterol monooxygenase 1-1-like
21932502	101507785	uncharacterized protein LOC101507785
21939299	101499761	protein cornichon homolog 4-like
21949526	101500079	AT-hook motif nuclear-localized protein 7-like
21964246	101508107	65-kDa microtubule-associated protein 9-like
21971367	101508422	LOB domain-containing protein 19
21991340	101500615	LOB domain-containing protein 18

Appendix 6.1 Information about the 96 chickpea accessions used in this study. The first column indicates whether the accessions correspond to a *desi* (D), *kabuli* (K) or wild (W) line. % PF is the proportion of the total reads assigned to each genotype. The last five columns indicate the presence (Y) or lack (N/-) of some of the most significant polymorphism described in chapter 6: The insertion of ~750 bp in the *FTa1-a2* intergenic region, the *FTa2* deletions *type 1* and *type 2* and the allele of the SNP (G269T) described in *FTc* gene (only given in those accessions on which *FTc* could be completely sequenced).

Type	SampleID	% (PF)	Country	750 Indel	*type1*	*type2*	FTc complete	FTcSNP
D	ICC13599	0.8413	Iran	N	-	-	Y	G
W	Cr5-9	1.1977	Turkey	N	-	-	Y	G
K	ICC2482	0.7964	Iran	N	-	-	Y	G
D	ICC2737	1.2643	Iran	N	-	-	Y	G
D	ICC2990	1.1682	Iran	N	-	-	Y	G
D	ICC3391	1.0974	Iran	N	-	-	Y	G
D	ICC3512	1.0289	Iran	N	-	-	Y	G
K	ICC9434	1.4676	Iran	N	-	-	Y	G
W	PI489777	0.6459	Turkey	N	-	-	Y	G
D	ICC12155	1.0977	Bangladesh	Y	-	-	Y	T
D	ICC12654	0.73	Ethiopia	N	-	-	Y	T
D	ICC4872	1.4772	India	N	-	-	Y	T
K	ICC7315	0.4882	Iran	N	-	-	Y	T
K	ICC9137	0.8184	Iran	N	-	-	Y	T
D	ICC9712	0.9295	Afghanistan	N	-	-	Y	T
D	ICC6306	0.2967	Russia & CISs	N	-	-	Y	T
D	ICC1180	0.578	India	Y	-	-	Y	T
D	ICC16374	0.7565	Malawi	Y	-	-	Y	T
D	ICC4958	0.8057	India	Y	-	-	Y	T
D	ICC8318	1.5205	India	Y	-	-	Y	T
K	ICCL81001	0.3431	Spain	Y	-	-	Y	T
K	ICC5878	1.3957	India	Y	-	-	Y	T
D	ICC5434	1.3618	India	N	-	Y	N	-
K	CA2156	1.2071	Spain	N	-	-	N	-
D	ICC1052	0.8247	Pakistan	N	-	-	N	-
K	ICC10755	1.1205	Turkey	N	-	-	N	-
D	ICC12824	0.7004	Ethiopia	N	-	-	N	-
D	ICC12866	0.5652	Ethiopia	N	-	-	N	-
D	ICC14051	0.3149	Ethiopia	N	-	-	N	-

Type	SampleID	% (PF)	Country	750 Indel	*type1*	*type2*	FTc complete	FTcSNP
K	ICC15435	0.9321	Morocco	N	-	-	N	-
D	ICC15567	0.0315	India	N	-	-	N	-
K	ICC15697	0.4712	Syrian Arab Rep.	N	-	-	N	-
K	ICC15802	1.0267	Syrian Arab Rep.	N	-	-	N	-
D	ICC2507	1.0511	Iran	N	-	-	N	-
D	ICC2884	1.3773	Iran	N	-	-	N	-
D	ICC3631	0.0936	Iran	N	-	-	N	-
D	ICC4182	1.2721	Iran	N	-	-	N	-
D	ICC4363	1.0912	Iran	N	-	-	N	-
D	ICC4463	1.3698	Iran	N	-	-	N	-
D	ICC4814	0.8951	Iran	N	-	-	N	-
D	ICC5504	0.1666	Mexico	N	-	-	N	-
D	ICC7184	0.9735	Turkey	N	-	-	N	-
K	ICC7308	1.1399	Peru	N	-	-	N	-
K	ICC7571	0.8254	Israel	N	-	-	N	-
K	ICC8151	0.9535	USA	N	-	-	N	-
D	ICC8200	0.5093	Iran	N	-	-	N	-
K	ICC8261	2.0252	Turkey	N	-	-	N	-
D	ICC8515	1.3361	Greece	N	-	-	N	-
D	ICC8522	1.3138	Italy	N	-	-	N	-
K	ICC8855	1.2954	Afghanistan	N	-	-	N	-
K	ILC3279	1.0036	Russia & CISs	N	-	-	N	-
D	ICC14098	0.405	Ethiopia	N	-	-	N	-
D	ICC9590	1.4104	Egypt	N	-	-	N	-
K	ICCV2	1.389	India	-	-	Y	N	-
D	WR315	0.6248	India	-	-	Y	N	-
D	JG62	0.8138	Spain	N	-	-	N	-
D	ICC4991	0.7738	India	Y	-	-	N	-
D	ICC5613	1.1651	India	Y	-	-	N	-
D	ICC7413	0.912	India	Y	-	-	N	-
K	ICC1161	0.9466	Pakistan	-	Y	-	N	-
D	ICC11664	1.2942	India	-	Y	-	N	-
D	ICC1923	0.9245	India	-	Y	-	N	-
D	ICC2210	1.4243	Algeria	-	Y	-	N	-
K	ICC4841	1.0319	Morocco	-	Y	-	N	-
K	ICC4973	0.8855	India	-	Y	-	N	-
D	ICC6802	1.4466	Iran	-	Y	-	N	-

Type	SampleID	% (PF)	Country	750 Indel	*type1*	*type2*	FTc complete	FTcSNP
D	ICC9586	0.8576	India	-	Y	-	N	-
D	ICC1194	0.7515	India	-	Y	-	N	-
D	ICC2242	0.8019	India	-	Y	-	N	-
D	ICC10393	0.624	India	-	-	Y	N	-
D	ICC10399	0.5715	India	-	-	Y	N	-
D	ICC12916	0.6954	India	-	-	Y	N	-
K	ICC12968	1.2846	India	-	-	Y	N	-
D	ICC13219	1.0465	Iran	-	-	Y	N	-
D	ICC1356	0.9305	India	-	-	Y	N	-
D	ICC14799	1.3964	India	-	-	Y	N	-
D	ICC15606	1.5501	India	-	-	Y	N	-
D	ICC15614	1.2316	Tanzania	-	-	Y	N	-
D	ICC15618	1.3348	India	-	-	Y	N	-
D	ICC15996	1.1629	India	-	-	Y	N	-
D	ICC16207	1.1299	Myanmar	-	-	Y	N	-
D	ICC16915	1.1681	India	-	-	Y	N	-
D	ICC1710	1.2785	India	-	-	Y	N	-
D	ICC2072	1.3593	India	-	-	Y	N	-
D	ICC2580	0.8036	Iran	-	-	Y	N	-
D	ICC440	1.1856	India	-	-	Y	N	-
D	ICC4533	1.184	India	-	-	Y	N	-
D	ICC5135	0.8483	India	-	-	Y	N	-
D	ICC6537	0.8462	Iran	-	-	Y	N	-
D	ICC6579	1.8071	Iran	-	-	Y	N	-
D	ICC6811	1.0412	Iran	-	-	Y	N	-
D	ICC6816	1.8146	Iran	-	-	Y	N	-
D	ICC95	0.9806	India	-	-	Y	N	-
D	ICC1398	1.5708	India	-	-	Y	N	-
D	ICC14669	1.5809	India	-	-	Y	N	-
D	ICC4639	0.961	India	-	Y	-	N	-

Appendix 6.2 Mean, Standard deviation (S) and standard error (Err) values (in days) obtained for flowering time (DTF) in the 96 chickpea accessions described in appendix 6.1, grown in four different environments; in a phytotron at Hobart (Tasmania) during 2015 and 2016, under either long days (LD) or Short days (SD); and in field at the Turretfield Research Station (South Australia).

Genotype	2016						2015			Turretfield	
	Mean LD	S (LD)	Err (LD)	Mean SD	S (SD)	Err (SD)	Mean LD	S	Err	DTF	S
ICC12968	26.8	3.0	1.5	34.5	3.5	1.8	25.0	2.0	1.2	69.0	1.2
ICC7413	27.8	2.2	1.1	44.7	5.0	2.9	23.0	1.7	1.0	84.0	3.4
ICC12654	24	0.0	0.0	31.8	1.5	0.8	20.7	0.6	0.3	87.3	2.4
ICC12824	26.3	1.9	0.9	34	4.2	2.1	21.0	1.0	0.6	87.3	3.7
ICC15435	33	1.7	1.0	57.3	4.5	2.6	25.7	3.5	2.0	87.3	5.2
ICC12155	28.8	1.3	0.6	50	14.0	7.0	23.7	1.5	0.9	87.7	3.6
ICC8318	28.5	3.0	1.5	41.3	8.8	4.4	23.7	1.5	0.9	87.7	3.4
ICC8522	27	1.4	0.7	38.3	3.1	1.8	21.7	0.6	0.3	87.7	3.8
ICC7308	31	1.4	0.7	51.3	13.6	6.8	26	2.0	1.2	88.7	3.5
ICC13219	31.5	2.4	1.2	41.75	2.9	1.4	26.0	0.0	0.0	88.7	3.9
ICC15614	29.8	2.9	1.4	48.8	4.5	2.3	23.0	0.0	0.0	89.3	4.8
ICC16374	29.8	1.0	0.5	53	14	8.1	22.7	0.6	0.3	89.7	5.0
ICC2507	25.5	2.5	1.3	40	4.4	2.5	22.0	0.0	0.0	89.7	2.5
ICC14098	25.3	1.5	0.8	38.7	9.3	5.4	20.3	1.5	0.9	90.0	3.5
ICC14669	31.5	1.9	1.0	48	5.7	2.8	26.0	1.7	1.0	90.0	3.9
ICC16915	31.5	0.6	0.3	48	2	1.0	24.7	2.5	1.5	90.0	4.8
ICC4363	24.8	0.5	0.3	56.25	12.9	6.4	20.7	0.6	0.3	91.0	2.9
ICC1356	33.7	1.2	0.6	47.5	6.1	3.1	27.0	2.0	1.2	91.0	4.7
ICC2580	30.3	1.5	0.8	48.25	2.4	1.2	25.3	1.5	0.9	91.0	3.5
ICC14051	25.3	1.0	0.5	35.5	3.4	1.7	21.7	0.6	0.3	92.0	2.5
ICC5878	30.5	1.9	1.0	57.8	11.8	5.9	23.3	1.2	0.7	92.0	5.1
ICC15996	33	1.4	0.7	53.3	8.7	4.3	27.7	1.2	0.7	92.0	3.8
ICC12866	25.8	1.0	0.5	34.7	2.1	1.0	22.3	3.2	1.9	93.0	2.4
ICC6816	33.8	0.5	0.3	62	4.8	2.4	26.3	1.2	0.7	93.0	5.3
ICC95	31	0.0	0.0	54	7.0	3.5	25.7	1.2	0.7	93.0	3.7
ICC15618	33.3	1.7	0.9	39.5	1.3	0.6	25	1.0	0.6	93.7	5.8
ICC4533	32.0	1.2	0.6	60.5	6.6	3.3	27.0	1.7	1.0	94.0	3.5
ICC5434	34.5	2.6	1.3	63.3	5.4	2.7	28.0	1.0	0.6	94.0	4.6
ICC3631	29.8	2.6	1.3	46	3.4	1.7	22.7	1.2	0.7	94.7	5.0
ICC4991	30.5	1.9	1.0	53.5	10.6	5.3	23.0	1.7	1.0	95.0	5.3
ICC6579	31.5	0.6	0.3	59	4.5	2.3	26.0	1.4	1.0	95.0	3.9
ICC8151	30.7	3.5	2.0	75.5	9.2	5.3	28.7	5.0	2.9	95.7	1.4

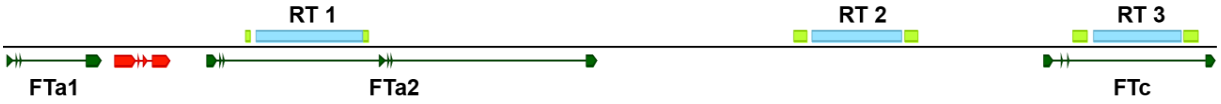
Appendix 6.2 Continued

Genotype	2016						2015			Turretfield	
	Mean LD	S (LD)	Err (LD)	Mean SD	S (SD)	Err (SD)	Mean LD	S	Err	DTF	S
ICC1398	30	2.0	1.0	37	1.8	0.9	23.7	1.5	0.9	96.0	4.5
ICC2482	26.8	3.9	1.9	50.5	17.7	8.9	21.5	0.7	0.4	96.7	3.7
ICC2737	29	3.7	1.9	45.5	8.7	4.3	22.7	1.2	0.7	97.0	4.5
ICC5613	32.3	1.0	0.5	53.3	4.6	2.3	28.7	4.5	2.6	97.0	2.5
ICC14799	32.3	1.3	0.6	59.3	8.1	4.0	27.3	1.5	0.9	97.0	3.5
ICC7315	28	1.2	0.6	46.7	7.1	4.1	23.0	1.7	1.0	97.7	3.5
ICC9712	25.8	3.0	1.5	63	7.9	4.6	21.0	1.0	0.6	97.7	3.4
ICC2884	28.3	1.0	0.5	50.3	16	9.5	22.3	0.6	0.3	98.0	4.2
ICC8200	28.3	2.1	1.0	63.3	1.5	0.9	26.7	8.1	4.7	98.3	1.1
ICC3391	26	1.4	0.7	54.5	9.9	5.0	22.0	0.0	0.0	98.7	2.8
ICC4182	27.5	0.6	0.3	42.3	2.5	1.3	23.0	1.4	1.0	98.7	3.2
ICC9590	26	0.8	0.4	32.8	2.1	1.0	21.7	0.6	0.3	99.0	3.1
ICC10755	31.8	2.5	1.3	60	6.9	4.0	26.3	3.8	2.2	99.3	3.8
ICC15802	27.3	1.7	0.9	60.7	7.5	4.3	22.3	1.2	0.7	99.3	3.5
ICC4463	27.8	3.2	1.6	49	14.1	10.0	21.3	1.2	0.7	99.3	4.6
ICC10399	31.3	0.5	0.3	55.8	4.3	2.1	25.5	0.7	0.4	99.3	4.1
ICC2990	29	1.7	1.0	45	-	-	25.0	5.2	3.0	99.7	2.8
ICC4814	27.8	2.8	1.4	47.7	7.5	4.3	22.0	0.0	0.0	99.7	4.1
ICC1180	30.3	1.0	0.5	61.3	5.6	2.8	24.3	2.5	1.5	100.3	4.2
ICC8261	30.8	1.5	0.8	121.5	9.2	6.5	26.3	5.1	3.0	100.3	3.1
ICC1052	28	3.6	1.8	41.3	5.5	3.2	21.5	0.7	0.5	100.7	4.6
ICC4639	34.0	0.8	0.4	67.5	1.9	1.0	28.0	1.7	1.0	100.7	4.2
ICC13599	28.3	0.6	0.3	95.5	2.1	1.5	25	3.6	2.1	101.0	2.4
ICC3512	26.5	0.6	0.3	38.7	6.5	3.8	23.3	3.2	1.9	101.0	2.3
ICC8855	25.8	1.7	0.9	56.3	15.9	7.9	21.0	1.0	0.6	101.3	3.4
ICC4973	29	0.8	0.4	62.3	2.1	1.0	25.0	2.0	1.2	101.7	2.8
ICC1194	33.5	0.6	0.3	52.8	6.7	3.4	26.3	1.5	0.9	102.3	5.1
ICC6802	34	0.8	0.4	65.5	0.7	0.5	29.0	0.0	0.0	103.0	3.5
ICC16207	34.0	0.0	0.0	75	5.7	2.9	28.3	1.2	0.7	103.0	4.0
ICC2072	31.8	1.5	0.8	57.3	5.1	2.6	27.3	0.6	0.3	103.3	3.1
ICC9137	32.3	1.5	0.9	68.7	1.5	0.9	27.0	3.5	2.0	104.0	3.8
ICC15606	32	2.0	1.0	52.5	0.6	0.3	26.7	0.6	0.3	104.0	3.8
ICC9434	29.3	2.1	1.2	113	7.5	4.4	28.0	-	-	105.0	0.9
ICC12916	32	0.0	0.0	68.5	4.4	2.2	27.7	1.5	0.9	105.0	3.1
ICC1710	32.5	0.6	0.3	64.3	1.5	0.8	28.0	1.4	1.0	105.0	3.2
ICC11664	34.3	0.5	0.3	66.3	0.5	0.3	26.7	0.6	0.3	106.0	5.3
ICC2210	32.8	0.5	0.3	71.0	2.4	1.2	28.7	0.6	0.3	106.0	2.9

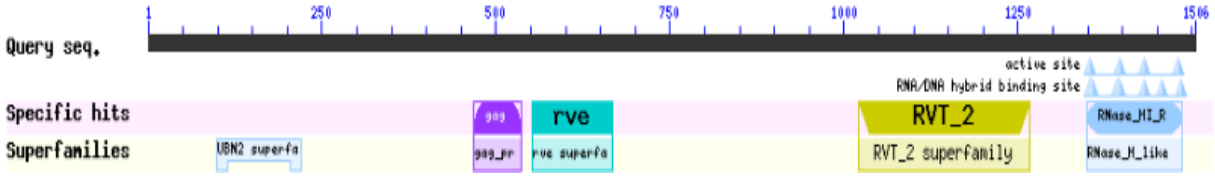
Appendix 6.2 Continued

Genotype	2016						2015			Turretfield	
	Mean LD	S (LD)	Err (LD)	Mean SD	S (SD)	Err (SD)	Mean LD	S	Err	DTF	S
ICC7571	31	5	2.9	68	2.8	1.6	29.3	7.5	4.3	106.3	1.2
ICC15697	24.7	1.5	0.9	41.5	3.5	2.5	37.0	5.7	4.0	107.0	8.7
ICC4841	31.0	0.8	0.4	65.3	3.8	1.9	28.0	1.4	1.0	107.0	2.1
ICC6811	32.5	1.0	0.5	66.3	1.5	0.8	27.0	1.0	0.6	107.0	3.9
ICC9586	35.3	2.1	1.0	76.8	1.7	0.9	30	1.7	1.0	108.0	3.7
ICC6306	33.3	1.0	0.5	64	1.7	0.9	27.0	1.7	1.0	109.0	4.4
ICC1161	33.8	1.0	0.5	108	4.2	3.0	29.0	-	-	109.0	3.4
ICC1923	32.5	1.0	0.5	67	6.1	3.0	27.0	2.0	1.2	109.0	3.9
ICC10393	32	1.2	0.6	53.3	2.5	1.5	25.7	1.5	0.9	109.0	4.5
ICC15567	31	2.2	1.1	68.3	6.7	3.4	28.0	1.0	0.6	110.0	2.1
ICC5504	27	0.8	0.4	80.0	1.4	0.7	22.7	0.6	0.3	110.0	3.0
ICC8515	30.8	1.0	0.5	64.7	8.5	4.9	25.3	0.6	0.3	110.0	3.8
ICC2242	33.5	0.6	0.3	79.7	4.0	2.0	27.7	1.2	0.7	110.0	4.1
ICC7184	53.3	1.3	0.6	93	1.4	1.0	44.0	0.0	0.0	112.0	6.6
ICC6537	31.5	1.0	0.5	69.3	7.6	3.8	27.7	1.2	0.7	118.0	2.7
ICC4872	27.5	0.7	0.5	64	NA	NA	27	2.8	2.0	118.7	0.4
ICC440	33.3	1.3	0.6	59.5	9.2	6.5	28.3	1.2	0.7	119.0	3.5
ICC5135	35.3	3.3	1.7	70.3	4.5	2.3	28.3	1.2	0.7	120.0	4.9
CA2156	34.7	0.6	0.3	160	1.4	1	-	-	-	-	-
Cr5-9	30.5	0.6	0.3	144.7	2.1	1.2	-	-	-	-	-
ICC4958	31.5	0.6	0.3	54.3	10.4	5.2	-	-	-	-	-
ICCL81001	32.5	0.6	0.3	36.5	1.9	1.0	-	-	-	-	-
ILC3279	37.8	9.2	4.6	105	32.3	16.2	-	-	-	-	-
JG62	33.3	1.0	0.5	153.7	12.1	7.0	-	-	-	-	-
PI489777	30	1.2	0.6	167	2.9	1.5	-	-	-	-	-
ICCV2	27.3	1.5	0.8	37	5.2	3.0	-	-	-	-	-
WR315	32.8	0.5	0.3	45.3	5.1	2.6	-	-	-	-	-

Appendix 6.3 Diagram of the FT cluster, indicating the position of the three retrotransposons (RT 1 to 3) found within the region. Further information regarding each of them can be found in appendixes 6.3.1 to 6.3.3.

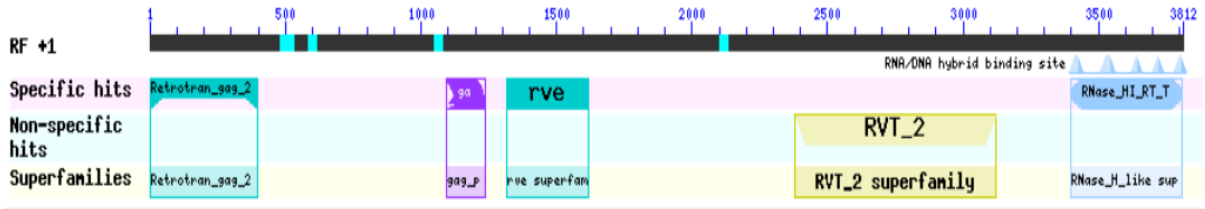


Appendix 6.3.1 Schematic diagram of retrotransposon 1, showing the relative position, name, accession number and E-value of the different domains found within its sequence.



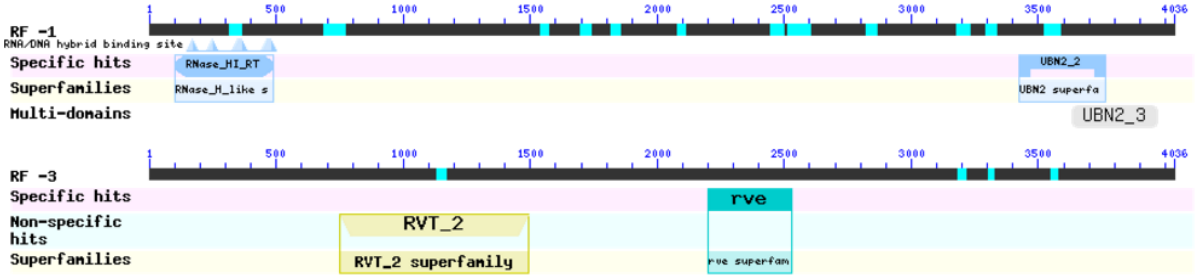
Name	Accession	Description	E-value
RVT_2	pfam07727	Reverse transcriptase (RNA-dependent DNA polymerase)	2.87E-121
RNase_HI_RT_Ty1	cd09272	Ribonuclease H	5.78E-78
rve	pfam00665	Integrase core domain	2.12E-18
gag_pre-integrs	pfam13976	AG-pre-integrase domain	3.21E-15
UBN2 super family	cl15874	gag-polypeptide of LTR copia-type	4.43E-07
UBN2_3	pfam14244	gag-polypeptide of LTR copia-type	3.60E-46

Appendix 6.3.2 Schematic diagram of retrotransposon 2, showing the relative position, name, accession number and E-value of the different domains found within its sequence.



Name	Accession	Description	E-value
RVT_2	pfam07727	Reverse transcriptase (RNA-dependent DNA polymerase)	1.75E-75
RNase_HI_RT_Ty1	cd09272	Ribonuclease H	6.33E-37
Retrotran_gag_2	pfam14223	gag-polypeptide of LTR copia-type	1.42E-21
rve	pfam00665	Integrase core domain	1.74E-09
gag_pre-integrs	pfam13976	GAG-pre-integrase domain	2.31E-03

Appendix 6.3.3 Schematic diagram of retrotransposon 3, showing the relative position, name, accession number and E-value of the different domains found within its sequence.



Name	Accession	Description	E-value
RNase_HI_RT_Ty1	cd09272	Ribonuclease H (RNase H)	2.02E-25
UBN2_2	pfam14227	gag-polypeptide of LTR copia-type	5.23E-09
UBN2_3	pfam14244	gag-polypeptide of LTR copia-type	8.15E-05
RVT_2	pfam07727	Reverse transcriptase (RNA-dependent DNA polymerase)	3.43E-55
rve	pfam00665	Integrase core domain	8.76E-08

Appendix 6.4 Accession number of the 109 chickpea lines screened at the University of Saskatoon (Canada) and their flowering phenotype under long days (LD) and short days (SD) with standard error (SE). The last column indicates the presence (Y) or deletion (N) of the FTa2 gene in each particular line.

Name	LD	SE	SD	SE	FTa2		Name	LD	SE	SD	SE	FTa2
ICC12004	25.3	4.6	72.3	4.9	Y		FLIP83-7C	49.3	4.2	91.0	6.5	Y
CDC Anna	39.8	7.0	99.0	6.3	Y		FLIP84-92C	52.5	3.5	109.2	5.4	Y
CDC Chichi	39.5	7.6	97.5	3.7	Y		FLIP86-5C	47.4	4.0	86.5	4.8	Y
CDC Chico	38.1	5.5	97.5	5.2	Y		FLIP86-6C	49.3	5.2	100.5	4.4	Y
CDC Xena	35.5	6.4	83.2	2.6	Y		FLIP85-1C	50.1	4.4	93.7	3.3	Y
CDC Luna	45.1	5.8	99.5	3.7	Y		FLIP85-17C	48.7	4.8	90.2	2.9	Y
95168-64	40.7	5.2	92.2	5.5	Y		FLIP87-45C	42.0	6.3	85.7	3.6	Y
95177-47	46.6	5.6	92.3	2.5	Y		FLIP87-8C	35.6	6.6	64.3	5.8	Y
CDC Frontier	43.9	5.7	106.8	9.5	Y		FLIP88-85C	45.9	5.2	97.3	3.1	Y
Amit	53.7	6.4	109.2	6.1	Y		FLIP90-96C	56.1	4.9	116.3	8.2	Y
CDC Ebony	36.0	3.9	84.7	4.9	Y		FLIP91-77C	44.3	6.5	75.3	6.0	Y
242-2	36.7	2.9	86.7	8.3	Y		FLIP93-93	47.1	5.7	88.7	5.7	Y
97-Indian2-112	38.1	4.4	100.8	4.8	Y		FLIP93-146C	56.7	4.5	96.2	10.8	Y
CDC Vanguard	35.9	3.7	76.2	3.7	Y		FLIP93-58C	45.7	6.3	94.0	3.5	Y
FLIP97-45C	37.7	3.5	73.7	7.5	Y		FLIP97-263C	39.0	7.8	53.5	2.4	Y
316B-42	37.4	2.6	78.2	4.9	Y		FLIP97-281C	38.6	7.7	74.7	8.6	Y
328S-8	37.7	3.0	86.7	3.0	Y		ILC 72	53.4	7.3	106.3	7.3	Y
FLIP98-135C	32.5	1.7	63.0	4.5	Y		ILC 195	50.7	7.6	95.5	3.9	Y
FLIP95-56C	41.1	4.4	83.0	4.6	Y		ILC 2555	50.4	7.4	105.3	5.9	Y
425-14	32.0	0.9	78.8	6.2	Y		ILC 3279	59.6	6.1	103.0	3.5	Y
438-29	33.0	1.4	84.2	2.0	Y		ILC482	45.4	9.1	76.7	7.1	Y
CIABN-99PL27119	48.7	2.1	99.8	2.0	Y		1041-3	50.3	8.1	88.0	2.0	Y
CDC Orion	34.6	1.7	71.5	4.3	Y		CDC Consul (603-3)	55.1	8.1	90.0	11.2	Y
603-3	48.3	2.8	90.5	5.9	Y		713-13	47.6	9.7	94.8	6.0	Y

Appendix 6.4 Continued

Name	LD	SE	SD	SE	FTa2		Name	LD	SE	SD	SE	FTa2
701-6	40.6	2.4	87.8	2.5	Y		CA05-75-45	50.4	9.0	99.7	3.7	Y
889-8	48.4	2.9	94.5	3.4	Y		1401-1	48.3	9.7	88.2	3.8	Y
CDC Corinne	55.7	2.9	87.5	4.0	Y		AB06-160-4	47.4	10.1	92.2	1.6	Y
1041-3	45.3	2.0	94.2	4.4	Y		1349-1	45.0	12.0	49.3	3.3	Y
1044-6	49.4	6.0	79.0	7.2	Y		AB06-106-2	48.2	12.1	87.2	6.9	Y
1045-1	43.1	3.1	96.2	3.5	Y		1460-2	52.6	10.4	82.0	10.6	Y
GPE094	52.8	3.1	88.0	7.4	Y		ICCV 96029	22.9	0.5	24.9	0.4	Y
Y9563-028	33.3	4.2	67.8	4.1	Y		95-NN-12	55.7	1.4	100.5	3.2	Y
CA05-73-6	55.6	1.8	109.0	6.9	Y		FLIP81-293C	42.6	4.1	96.5	3.5	Y
CA05-75-16	41.3	3.9	101.0	3.6	Y		FLIP82-150C	46.7	3.1	94.5	4.6	Y
BS1-D-15	36.7	6.5	86.7	6.8	N		561aS-18	42.0	2.0	86.0	2.7	N
DH45-1	36.7	6.3	81.3	1.4	N		612-4	44.9	3.3	82.3	3.5	N
Myles	31.4	5.0	72.2	3.5	N		FLIP81-71C	44.9	2.9	88.3	2.6	N
CDC Cabri	30.7	4.5	70.3	2.8	N		FLIP84-48C	49.8	3.8	89.0	2.8	N
CDC Desiray	30.6	4.0	80.2	4.2	N		FLIP84-188C	47.3	3.9	85.5	1.7	N
CDC Nika	28.4	3.5	70.0	4.9	N		FLIP97-137C	50.0	5.3	87.7	2.8	N
CDC Verano	38.2	7.4	90.7	4.8	N		FLIP97-503C	47.6	6.4	91.7	5.0	N
CDC Yuma	36.9	4.6	87.0	2.9	N		FLIP97-530C	45.0	6.9	83.0	2.1	N
FLIP95-48C	34.8	3.4	74.0	4.1	N		FLIP97-677C	49.7	6.5	89.0	3.3	N
S95420	36.7	3.4	85.7	2.0	N		FLIP97-706C	46.3	7.1	79.0	2.1	N
CDC Jade	40.4	3.5	81.7	3.4	N		FLIP98-121C	48.4	7.9	82.8	7.1	N
FLIP97-101C	43.1	3.6	99.5	2.2	N		ILC588	45.4	7.6	66.5	5.6	N
381T-4	43.4	3.8	108.7	5.5	N		ILC 484	49.7	7.6	88.8	1.6	N
418-59	44.7	2.5	77.3	2.3	N		Elixir	52.3	7.6	93.2	3.2	N
441-34	40.4	1.6	76.7	2.9	N		551-1 *	45.9	9.3	76.2	3.8	N
CA2969	39.7	0.9	71.5	3.1	N		1173-1	48.1	9.3	85.3	4.6	N
492-3	41.7	1.8	83.5	4.4	N		AB06-156-2	50.1	9.6	90.7	3.6	N

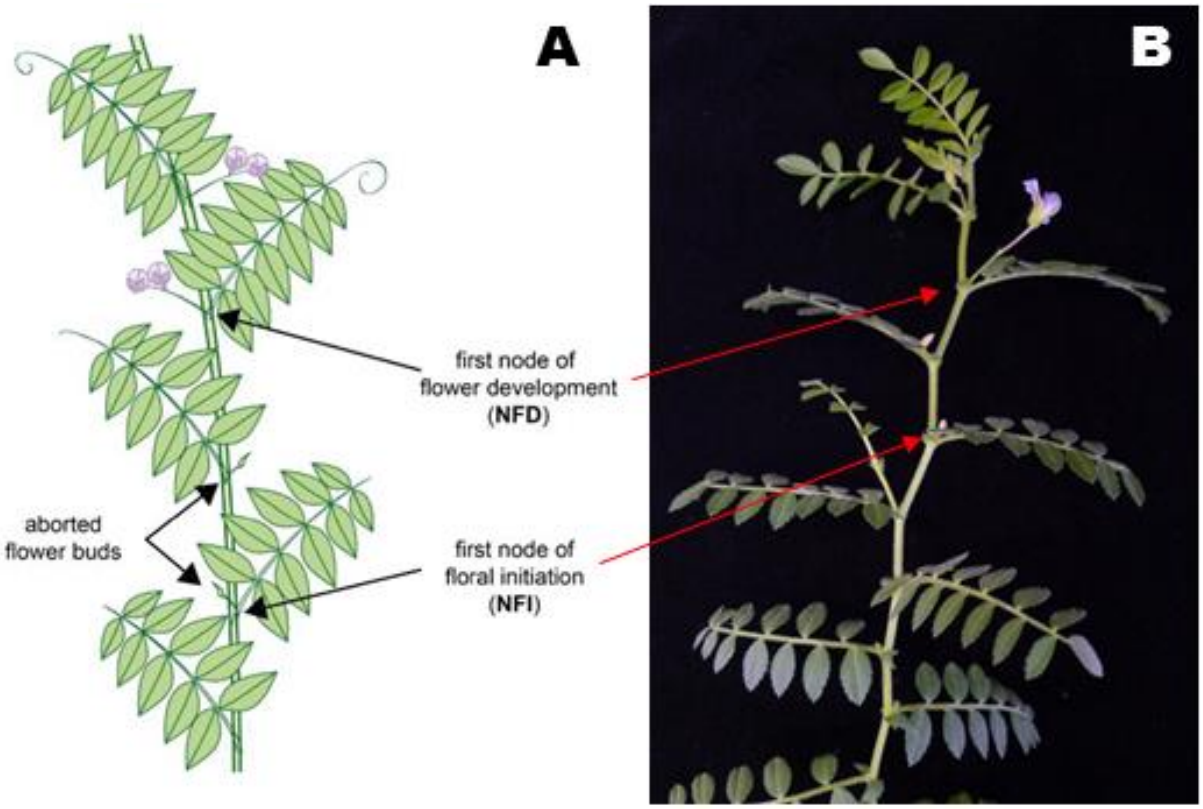
Appendix 6.4 Continued

Name	LD	SE	SD	SE	FTa2		Name	LD	SE	SD	SE	FTa2
CDC Leader	40.6	0.6	75.6	5.9	N		X05TH47-3	55.0	9.2	80.7	3.1	N
494-4	43.9	1.3	80.5	7.5	N		X05TH20-2	51.1	9.9	89.0	3.0	N
512-51	40.9	1.0	67.5	5.4	N		553-1	41.9	1.3	74.3	4.1	N
551-1	40.0	1.7	71.0	4.2	N							

Appendix 6.5 Haplotypes detected by PopART software and used to construct the median-joining network displayed on figure 6.14.

Single accession Haplotypes	Haplotypes with multiple accessions			
	Haplotype name	**All accessions included in the haplotype**		
ICC12824	**1052**	1052	2884	4182
ICC15435		8855	4814	2507
ICC15618	**12155**	12155	4991	1180
ICC2482	**V2**	V2	14669	12698
ICC2737		16915	15614	
ICC3631	**2580**	2580	15996	
ICC4463	**440**	440	16374	4958
ICC4973		5135	4872	8318
ICC6306		6811	12654	81001
ICC7184		WR315	9712	5878
ICC8261		1398	JG62	7413
ICC8515		4533	7315	5613
ICC9434		6537	15697	1923
ICC95		16207	3279	4639
ICC9590		6579	9137	9586
ICC10755		1710	CA2156	4841
ICC4363		1356	8200	2242
ICC15567		13219	8151	2210
ICC2072		12916	8522	11664
ICC3512		6816	14098	1161
		5434	10393	
	1194	1194	6802	
	10399	10399	15606	14799
	2990	2990	13599	3391
	7308	7308	5504	
	15802	15802	7571	
	12866	12866	14051	
	PI489777	PI489777	Cr5-9	

Appendix 7.1 A) Schematic of chickpea plant illustrating the node of flower initiation (NFI, or node bearing the first aborted flower bud), node of flower development (NFD, node bearing the first open flower), and aborted flower buds. Adapted from (Rajandran et al). **B)** Picture of chickpea plant showing NFI and NFD.



Appendix 7.2 Analysis of statistical significance between values obtained for the variable "days from emergence to first open flower" during 2015 season in 8 chickpea genotypes and 4 conditions: Long days (LD), short days (SD), vernalized (V) and non-vernalized (NV).

		Cr5-9			ICC81001		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	29.600*	8.442	.001	1.286	3.988	.747
	SDNV	-62.733*	5.628	.000	-24.800*	4.393	.000
	SDV	23.600*	6.448	.000	2.000	6.092	.743
LDV	LDNV	-29.600*	8.442	.001	-1.286	3.988	.747
	SDNV	-92.333*	8.898	.000	-26.086*	4.512	.000
	SDV	-6.000	9.438	.526	.714	6.179	.908
SDNV	LDNV	62.733*	5.628	.000	24.800*	4.393	.000
	LDV	92.333*	8.898	.000	26.086*	4.512	.000
	SDV	86.333*	7.035	.000	26.800*	6.448	.000
SDV	LDNV	-23.600*	6.448	.000	-2.000	6.092	.743
	LDV	6.000	9.438	.526	-.714	6.179	.908
	SDNV	-86.333*	7.035	.000	-26.800*	6.448	.000
		ICC4958			**ICCV2**		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	5.043	3.191	.115	2.429	3.567	.497
	SDNV	-21.100*	3.655	.000	-11.514*	4.512	.011
	SDV	3.218	3.367	.340	3.055	3.613	.399
LDV	LDNV	-5.043	3.191	.115	-2.429	3.567	.497
	SDNV	-26.143*	3.415	.000	-13.943*	4.015	.001
	SDV	-1.825	3.105	.557	.626	2.968	.833
SDNV	LDNV	21.100*	3.655	.000	11.514*	4.512	.011
	LDV	26.143*	3.415	.000	13.943*	4.015	.001
	SDV	24.318*	3.581	.000	14.569*	4.055	.000
SDV	LDNV	-3.218	3.367	.340	-3.055	3.613	.399
	LDV	1.825	3.105	.557	-.626	2.968	.833
	SDNV	-24.318*	3.581	.000	-14.569*	4.055	.000

Appendix 7.2 Continued

		ILC3279			PI489777		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	45.308*	4.055	.000	32.651*	3.884	.000
	SDNV	-89.000*	5.170	.000	-59.971*	3.798	.000
	SDV	40.750*	4.393	.000	23.054*	3.988	.000
LDV	LDNV	-45.308*	4.055	.000	-32.651*	3.884	.000
	SDNV	-134.308*	4.406	.000	-92.622*	3.541	.000
	SDV	-4.558	3.463	.190	-9.597*	3.745	.011
SDNV	LDNV	89.000*	5.170	.000	59.971*	3.798	.000
	LDV	134.308*	4.406	.000	92.622*	3.541	.000
	SDV	129.750*	4.719	.000	83.025*	3.655	.000
SDV	LDNV	-40.750*	4.393	.000	-23.054*	3.988	.000
	LDV	4.558	3.463	.190	9.597*	3.745	.011
	SDNV	-129.750*	4.719	.000	-83.025*	3.655	.000
		JG62			WR315		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	83.078*	3.726	.000	2.583	3.398	.448
	SDNV	-22.286*	8.238	.007	-11.500*	4.062	.005
	SDV	74.623*	3.726	.000	.231	3.342	.945
LDV	LDNV	-83.078*	3.726	.000	-2.583	3.398	.448
	SDNV	-105.364*	8.049	.000	-14.083*	3.853	.000
	SDV	-8.455*	3.286	.011	-2.353	3.085	.447
SDNV	LDNV	22.286*	8.238	.007	11.500*	4.062	.005
	LDV	105.364*	8.049	.000	14.083*	3.853	.000
	SDV	96.909*	8.049	.000	11.731*	3.803	.002
SDV	LDNV	-74.623*	3.726	.000	-.231	3.342	.945
	LDV	8.455*	3.286	.011	2.353	3.085	.447
	SDNV	-96.909*	8.049	.000	-11.731*	3.803	.002

*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix 7.3 Analysis of statistical significance between values obtained for the variable "days from emergence to first open flower" during 2016 season in 8 chickpea genotypes and 4 conditions: Long days (LD), short days (SD), vernalized (V) and non-vernalized (NV).

		Cr5-9			ICC81001		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	44.062*	4.012	.000	2.067	4.580	.652
	SDNV	-96.326*	4.063	.000	-19.711*	4.134	.000
	SDV	36.896*	4.063	.000	-2.148	4.394	.625
LDV	LDNV	-44.062*	4.012	.000	-2.067	4.580	.652
	SDNV	-140.389*	3.889	.000	-21.778*	4.407	.000
	SDV	-7.167	3.889	.066	-4.214	4.652	.365
SDNV	LDNV	96.326*	4.063	.000	19.711*	4.134	.000
	LDV	140.389*	3.889	.000	21.778*	4.407	.000
	SDV	133.222*	3.941	.000	17.563*	4.214	.000
SDV	LDNV	-36.896*	4.063	.000	2.148	4.394	.625
	LDV	7.167	3.889	.066	4.214	4.652	.365
	SDNV	-133.222*	3.941	.000	-17.563*	4.214	.000
		ICC4958			**ICCV2**		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	-.583	7.633	.939	1.738	4.652	.709
	SDNV	-55.833*	7.633	.000	-12.133*	5.063	.017
	SDV	-25.083*	7.633	.001	-2.070	4.360	.635
LDV	LDNV	.583	7.633	.939	-1.738	4.652	.709
	SDNV	-55.250*	8.361	.000	-13.871*	4.896	.005
	SDV	-24.500*	8.361	.004	-3.808	4.165	.361
SDNV	LDNV	55.833*	7.633	.000	12.133*	5.063	.017
	LDV	55.250*	8.361	.000	13.871*	4.896	.005
	SDV	30.750*	8.361	.000	10.063*	4.620	.030
SDV	LDNV	25.083*	7.633	.001	2.070	4.360	.635
	LDV	24.500*	8.361	.004	3.808	4.165	.361
	SDNV	-30.750*	8.361	.000	-10.063*	4.620	.030

Appendix 7.3 Continued

		ILC3279			PI489777		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	19.600*	6.476	.003	16.926*	3.788	.000
	SDNV	-111.000*	5.827	.000	-118.600*	3.901	.000
	SDV	-5.000	6.476	.441	6.600	3.739	.078
LDV	LDNV	-19.600*	6.476	.003	-16.926*	3.788	.000
	SDNV	-130.600*	6.924	.000	-135.526*	3.948	.000
	SDV	-24.600*	7.478	.001	-10.326*	3.788	.007
SDNV	LDNV	111.000*	5.827	.000	118.600*	3.901	.000
	LDV	130.600*	6.924	.000	135.526*	3.948	.000
	SDV	106.000*	6.924	.000	125.200*	3.901	.000
SDV	LDNV	5.000	6.476	.441	-6.600	3.739	.078
	LDV	24.600*	7.478	.001	10.326*	3.788	.007
	SDNV	-106.000*	6.924	.000	-125.200*	3.901	.000
		JG62			WR315		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	47.941*	4.189	.000	3.121	3.999	.436
	SDNV	-82.337*	3.999	.000	-16.824*	3.999	.000
	SDV	31.731*	3.948	.000	-4.402	3.948	.265
LDV	LDNV	-47.941*	4.189	.000	-3.121	3.999	.436
	SDNV	-130.278*	4.134	.000	-19.944*	3.941	.000
	SDV	-16.211*	4.084	.000	-7.523	3.889	.054
SDNV	LDNV	82.337*	3.999	.000	16.824*	3.999	.000
	LDV	130.278*	4.134	.000	19.944*	3.941	.000
	SDV	114.067*	3.889	.000	12.421*	3.889	.002
SDV	LDNV	-31.731*	3.948	.000	4.402	3.948	.265
	LDV	16.211*	4.084	.000	7.523	3.889	.054
	SDNV	-114.067*	3.889	.000	-12.421*	3.889	.002

*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix 7.4. Analysis of statistical significance between values obtained for the variable "days from emergence to first pod" during 2015 season in 8 chickpea genotypes and 4 conditions: Long days (LD), short days (SD), vernalized (V) and non-vernalized (NV).

		Cr5-9			ICC81001		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	45.600*	8.856	.000	5.143	4.184	.221
	SDNV	.[b]	.	.	-30.600*	4.609	.000
	SDV	36.600*	6.764	.000	7.000	6.392	.275
LDV	LDNV	-45.600*	8.856	.000	-5.143	4.184	.221
	SDNV	.[b]	.	.	-35.743*	4.734	.000
	SDV	-9.000	9.902	.365	1.857	6.482	.775
SDNV	LDNV	.[c]	.	.	30.600*	4.609	.000
	LDV	.[c]	.	.	35.743*	4.734	.000
	SDV	.[c]	.	.	37.600*	6.764	.000
SDV	LDNV	-36.600*	6.764	.000	-7.000	6.392	.275
	LDV	9.000	9.902	.365	-1.857	6.482	.775
	SDNV	.[b]	.	.	-37.600*	6.764	.000
		ICC4958			**ICCV2**		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	29.200*	3.347	.000	3.143	3.742	.402
	SDNV	-30.900*	4.428	.000	-18.114*	4.734	.000
	SDV	27.518*	3.532	.000	2.619	3.845	.497
LDV	LDNV	-29.200*	3.347	.000	-3.143	3.742	.402
	SDNV	-60.100*	4.212	.000	-21.257*	4.212	.000
	SDV	-1.682	3.257	.606	-.524	3.181	.869
SDNV	LDNV	30.900*	4.428	.000	18.114*	4.734	.000
	LDV	60.100*	4.212	.000	21.257*	4.212	.000
	SDV	58.418*	4.361	.000	20.733*	4.303	.000
SDV	LDNV	-27.518*	3.532	.000	-2.619	3.845	.497
	LDV	1.682	3.257	.606	.524	3.181	.869
	SDNV	-58.418*	4.361	.000	-20.733*	4.303	.000

Appendix 7.4 Continued

		ILC3279			PI489777		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	56.583*	4.303	.000	41.778*	4.261	.000
	SDNV	.[b]	.	.	-43.333*	8.732	.000
	SDV	46.750*	4.609	.000	39.542*	4.366	.000
LDV	LDNV	-56.583*	4.303	.000	-41.778*	4.261	.000
	SDNV	.[b]	.	.	-85.111*	8.522	.000
	SDV	-9.833*	3.690	.008	-2.236	3.928	.570
SDNV	LDNV	.[c]	.	.	43.333*	8.732	.000
	LDV	.[c]	.	.	85.111*	8.522	.000
	SDV	.[c]	.	.	82.875*	8.575	.000
SDV	LDNV	-46.750*	4.609	.000	-39.542*	4.366	.000
	LDV	9.833*	3.690	.008	2.236	3.928	.570
	SDNV	.[b]	.	.	-82.875*	8.575	.000
		JG62			**WR315**		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	.[c]	.	.	9.444*	3.565	.009
	SDNV	.[b,c]	.	.	-14.389*	4.261	.001
	SDV	.[c]	.	.	6.085	3.506	.084
LDV	LDNV	.[b]	.	.	-9.444*	3.565	.009
	SDNV	.[b]	.	.	-23.833*	4.042	.000
	SDV	-9.727*	3.447	.005	-3.359	3.236	.301
SDNV	LDNV	.[b,c]	.	.	14.389*	4.261	.001
	LDV	.[c]	.	.	23.833*	4.042	.000
	SDV	.[c]	.	.	20.474*	3.990	.000
SDV	LDNV	.[b]	.	.	-6.085	3.506	.084
	LDV	9.727*	3.447	.005	3.359	3.236	.301
	SDNV	.[b]	.	.	-20.474*	3.990	.000

*. The mean difference is significant at the .05 level.

b. The level combination of factors in (J) is not observed.

c. The level combination of factors in (I) is not observed.

Appendix 7.5 Analysis of statistical significance between values obtained for the variable "days from emergence to first pod" during 2016 season in 8 chickpea genotypes and 4 conditions: Long days (LD), short days (SD), vernalized (V) and non-vernalized (NV).

		Cr5-9			ICC81001		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	48.385*	3.763	.000	1.833	4.295	.670
	SDNV	-104.618*	3.810	.000	-24.722*	3.877	.000
	SDV	40.160*	3.810	.000	-4.119	4.121	.318
LDV	LDNV	-48.385*	3.763	.000	-1.833	4.295	.670
	SDNV	-153.003*	3.647	.000	-26.556*	4.133	.000
	SDV	-8.225*	3.647	.025	-5.952	4.362	.173
SDNV	LDNV	104.618*	3.810	.000	24.722*	3.877	.000
	LDV	153.003*	3.647	.000	26.556*	4.133	.000
	SDV	144.778*	3.696	.000	20.603*	3.952	.000
SDV	LDNV	-40.160*	3.810	.000	4.119	4.121	.318
	LDV	8.225*	3.647	.025	5.952	4.362	.173
	SDNV	-144.778*	3.696	.000	-20.603*	3.952	.000
		ICC4958			ICCV2		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	-.083	7.158	.991	-.226	4.362	.959
	SDNV	-56.583*	7.158	.000	-15.883*	4.748	.001
	SDV	-31.083*	7.158	.000	-2.689	4.089	.511
LDV	LDNV	.083	7.158	.991	.226	4.362	.959
	SDNV	-56.500*	7.841	.000	-15.657*	4.591	.001
	SDV	-31.000*	7.841	.000	-2.462	3.906	.529
SDNV	LDNV	56.583*	7.158	.000	15.883*	4.748	.001
	LDV	56.500*	7.841	.000	15.657*	4.591	.001
	SDV	25.500*	7.841	.001	13.195*	4.332	.002
SDV	LDNV	31.083*	7.158	.000	2.689	4.089	.511
	LDV	31.000*	7.841	.000	2.462	3.906	.529
	SDNV	-25.500*	7.841	.001	-13.195*	4.332	.002

Appendix 7.5 Continued

		ILC3279			PI489777		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	19.400*	6.074	.002	36.134*	3.552	.000
	SDNV	-116.000*	5.465	.000	-115.858*	3.951	.000
	SDV	-1.400	6.074	.818	17.800*	3.507	.000
LDV	LDNV	-19.400*	6.074	.002	-36.134*	3.552	.000
	SDNV	-135.400*	6.493	.000	-151.992*	3.991	.000
	SDV	-20.800*	7.013	.003	-18.334*	3.552	.000
SDNV	LDNV	116.000*	5.465	.000	115.858*	3.951	.000
	LDV	135.400*	6.493	.000	151.992*	3.991	.000
	SDV	114.600*	6.493	.000	133.658*	3.951	.000
SDV	LDNV	1.400	6.074	.818	-17.800*	3.507	.000
	LDV	20.800*	7.013	.003	18.334*	3.552	.000
	SDNV	-114.600*	6.493	.000	-133.658*	3.951	.000
		JG62			WR315		
(I)	(J)	Mean Diff (I-J)	Std. Error	Sig.[b]	Mean Diff (I-J)	Std. Error	Sig.[b]
LDNV	LDV	47.167*	3.985	.000	.977	3.750	.795
	SDNV	-92.188*	3.921	.000	-20.967*	3.750	.000
	SDV	28.816*	3.763	.000	-9.096*	3.702	.014
LDV	LDNV	-47.167*	3.985	.000	-.977	3.750	.795
	SDNV	-139.354*	3.985	.000	-21.944*	3.696	.000
	SDV	-18.351*	3.830	.000	-10.073*	3.647	.006
SDNV	LDNV	92.188*	3.921	.000	20.967*	3.750	.000
	LDV	139.354*	3.985	.000	21.944*	3.696	.000
	SDV	121.003*	3.763	.000	11.871*	3.647	.001
SDV	LDNV	-28.816*	3.763	.000	9.096*	3.702	.014
	LDV	18.351*	3.830	.000	10.073*	3.647	.006
	SDNV	-121.003*	3.763	.000	-11.871*	3.647	.001

*. The mean difference is significant at the .05 level.